

1 **Demography, class-structure and kin selection in continuous-time models**

2

3 António M. M. Rodrigues^{1,2,*}

4

5 1. Department of Zoology, University of Cambridge, Downing Street, CB2 3EJ Cambridge,
6 United Kingdom.

7

8 2. Wolfson College, Barton Road, CB3 9BB Cambridge, United Kingdom.

9

10 * email: ammr3@cam.ac.uk

11

12 **Running head** -- Behavioural ecology and inclusive fitness

13

14

15

16

17

18

19

20

21

22

23

24

25 **Abstract** -- The Wright-Fisher infinite island model and the neighbour-modulated approach
26 to kin selection have enabled major advances in the understanding of social evolution in a
27 demographic context. Due to structural assumptions, however, some important evolutionary
28 problems are difficult to solve within the Wright-Fisher discrete-time framework. Although
29 these major constraints are relaxed in the Moran continuous-time framework, a formal
30 treatment of the mathematics of kin selection in continuous-time class-structured populations
31 is still lacking. Here, I employ the neighbour-modulated approach to formalise key features of
32 the kin selection theory in continuous-time infinite-island models. Next, I derive a general
33 form of Hamilton's rule to enable an inclusive fitness interpretation of social behaviours. I
34 consider class-structure at the group and individual level, and I focus on conditional and
35 unconditional phenotypes. I illustrate how the general theory can be applied to solve a wide
36 range of biological problems. Finally, I show how a simple extension of the framework
37 allows for the study of problems pertaining to the transmission of parental quality. I show that
38 while inheritance of parental quality may either promote or inhibit selection on conditional
39 helping behaviours, unconditional behaviours are invariant with respect to the fidelity of
40 inheritance.

41

42 **Keywords** -- class-structure, direct fitness, life-history, reproductive value, social evolution,
43 stochastic processes.

44

45

46

47

48

49

50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73

Introduction

Kin selection theory provides a general framework for studying the adaptive evolution of behaviours that affect not only the fitness of their bearers but also the fitness of their bearers' social partners [1-4]. Kin selection is elegantly encapsulated in Hamilton's rule, $-c + br > 0$, which provides the condition for the evolution of social traits [1,5]. A behaviour that inflicts a fitness cost c on the actor may nevertheless evolve if the fitness benefit b enjoyed by the recipient, depreciated by the relatedness r between actor and recipient, is sufficiently high. In simple scenarios, the canonical form of Hamilton's rule can be readily applied. More often, however, multiple ecological, demographic and genetic variables will influence the leading quantities in Hamilton's rule. Understanding social behaviour in such complex scenarios is a very active area of research and remains a major challenge for behavioural and evolutionary ecologists [4,6-9].

The inclusive-fitness and the neighbour-modulated approach to kin selection have been the two most widely used theoretical tools to study the evolution of social traits [1,6,10-13]. The inclusive fitness approach is actor-centric and focuses on how a trait expressed by a focal actor affects not only her own reproductive success but also the reproductive success of her social partners [12,13]. By contrast, the neighbour-modulated approach is recipient-centric, and it focuses on how a set of actors, composed by the social neighbours of the recipient, affect the reproductive success of the focal recipient [10,12,13]. While the formulation of these two approaches differs, they are mathematically equivalent, and therefore they yield the same predictions about the evolution of social traits.

74 Despite this equivalence, the neighbour-modulated approach has been particularly important
75 for solving kin selection problems that unfold in a rich demographic context [6,8,11,12,14-
76 26], where an inclusive-fitness approach is often harder to formulate. From the biological
77 details of the problem, one can define the neighbour-modulated fitness of a focal recipient,
78 and apply a well-studied mathematical optimisation method to derive the conditions under
79 which social behaviours evolve [6,10,11]. In principle, we can then conceptualise these
80 conditions in the form of Hamilton's rule (e.g. [15-18]). In practice, however, not all studies
81 obtain Hamilton's rule, and when Hamilton's rule is derived, it is often circumstantial and
82 inconsistent between studies. Thus a general formulation of Hamilton's rule that holds under
83 very general ecological and demographic conditions is still lacking.

84

85 The most widely used framework to model demography has been the Wright-Fisher infinite
86 island setting [27], in which populations are subdivided into several subpopulations connected
87 by dispersal [15,19-26,28-32]. In part, the success of the Wright-Fisher infinite-island
88 framework lies in being able to accommodate a considerable amount of ecological and
89 demographic detail and still deliver analytical solutions. However, the Wright-Fisher
90 framework also carries some structural assumptions that often limit the range of biological
91 problems that can be addressed (see Rodrigues and Kokko [9] for a review). Habitat
92 saturation is perhaps one of the most salient structural limitations of the Wright-Fisher
93 framework, as most natural populations evolve in unsaturated habitats where empty sites are
94 common.

95

96 A more recent breed of continuous-time models eases some of the most restrictive
97 assumptions of the Wright-Fisher framework. More specifically, continuous-time models
98 allow for the analysis of kin selection in unsaturated environments and account for ecological

99 feedbacks, while keeping the analytic nature of the Wright-Fisher framework (e.g. Alizon and
100 Taylor [33]; Wild *et al.* [34]; Rodrigues [35]). Alizon and Taylor, and Wild *et al.*, however,
101 have relied on the more heuristic inclusive-fitness approach to kin selection. In the limited
102 context of age-dependent social behaviour, I was able to derive Hamilton's rule using both the
103 inclusive-fitness and the neighbour-modulated approach, but I did not provide a formal
104 treatment of the relationship between stochastic processes, which underlie the continuous-
105 time models, and kin selection [35]. Therefore, a general formulation of the neighbour-
106 modulated approach to kin selection in stochastic continuous-time models as well as a formal
107 link between kin selection and the fundamental theory of stochastic processes remains
108 unclear.

109

110 Here, I develop a general framework linking together the formal theory of kin selection and
111 the fundamental theory of stochastic continuous-time models. First, I partition fitness into
112 different components according to the properties of the demographic structure. Second, I
113 establish a link between the fundamental theory of stochastic processes, kin selection and
114 demography. Finally, I use the neighbour-modulated approach to provide a general account of
115 kin selection in continuous-time models. Moreover, I consider between patch heterogeneity,
116 where patches differ in quality (e.g. [15,26]), but also within patch heterogeneity, where
117 individuals differ in quality (e.g. [17,31]). I consider behaviours that are expressed
118 conditionally on the quality of both actor and recipient, but also behaviours that are expressed
119 conditionally on the quality of patches, and behaviours that are expressed unconditionally.
120 For each case, I show how to derive Hamilton's rule using the neighbour-modulated
121 approach, and how to obtain an inclusive fitness interpretation of the behaviour. I then
122 illustrate how to apply the general framework to solve different biological problems,
123 including the evolution of cooperation, virulence, and age-dependent helping. Finally, I show

124 how a simple extension of this framework allows for the study of problems in which parental
125 and offspring quality are correlated.

126

127 **Kin selection: General principles**

128

129 I start off by providing a general formulation of fitness in class-structured populations
130 [6,10,11,17,31,36]. In class-structured populations, individuals belong to different classes, in
131 which class membership is associated with an individual's contribution to the evolutionary
132 process. Sex and age classes, for instance, are two of the most frequently observed class-
133 structures in natural populations. However, class-structures pertaining to other forms of
134 genetic or phenotypic differences among individuals are also common, such as class-structure
135 that emerges from differences in the nutritional, physiological, or environmental state of
136 individuals (e.g. [15,26,31]). In such general context, the neighbour-modulated fitness of a
137 random individual in the population is given by

138

$$139 \quad w = \sum_{\beta \in \Psi} u_{\beta} v_{\beta} W_{\beta}, \quad (1)$$

140

141 where: Ψ is the set of all possible classes; u_{β} is the frequency of individuals in the class β ; v_{β}
142 is the reproductive value of an individual in the class β ; and W_{β} is the neighbour-modulated
143 fitness of a random individual in the class β [6,10,11,31]. The fitness W_{β} of a random class- β
144 individual is given by the individual's genetic contribution to each class, denoted by $w_{\beta \rightarrow \eta}$,
145 weighted by the reproductive value of each recipient class relative to the reproductive value of
146 a class- β individual [6,10,31]. Thus, the fitness of a focal class- β individual is given by

147

$$148 \quad W_{\beta} = \frac{\sum_{\eta \in \Psi} w_{\beta \rightarrow \eta} v_{\eta}}{v_{\beta}}. \quad (2)$$

149

150 The set of class-specific reproductive successes, $w_{\beta \rightarrow \eta}$, form the elements of a square fitness
151 matrix, where each row is associated with a recipient class, while each column is associated
152 with a contributing class [6,10,17,31]. The fitness matrix is then given by

153

$$154 \quad \mathbf{A} = (w_{\beta \rightarrow \eta})_{\zeta \times \zeta}. \quad (3)$$

155

156 where ζ is the total number of classes in the population. The frequency and reproductive value
157 of individuals in each class can be directly calculated from the fitness matrix [6,10,17,31,36].

158 The frequency of individuals in each class is given by the right-eigenvector of matrix \mathbf{A} , while
159 the reproductive value of each individual is given by the left-eigenvector, with both
160 eigenvectors corresponding to the leading eigenvalue [10,17,31,36].

161

162 The frequency and reproductive value of individuals in each class can also be formulated as a
163 dynamical system (e.g. [37]). Let $\mathbf{u} = (u_1, u_2, \dots, u_\zeta)$ be the column vector whose elements
164 give the frequency of individuals in each class at any given time. The change over time in the
165 frequency of individuals in each class is described by a dynamical system given by

166

$$167 \quad \mathbf{u}' = (w_{\beta \rightarrow \eta})\mathbf{u}, \quad (4)$$

168

169 where \mathbf{u}' is the frequency of individuals in each class in the next time period. The changes in
170 reproductive value over time obey to a similar system of equations, which is given by

171

$$172 \quad \mathbf{v}' = \mathbf{v}(w_{\beta \rightarrow \eta}), \quad (5)$$

173

174 where $\mathbf{v} = (v_1, v_2, \dots, v_\zeta)$ is the row vector whose elements give the reproductive value of
175 each individual in each class. Like the right- and left-eigenvector results, the solution of these
176 systems of equations gives the asymptotic frequency and the reproductive value of individuals
177 in each class (e.g. [37]).

178

179 **Kin selection and demography**

180

181 My aim is to provide a general formulation of the fitness effects of a social trait z in a
182 demographic context assuming vanishingly small genetic variation in the population
183 [6,10,36]. I consider an infinite island model so that dispersal between patches is random and
184 uniformly distributed [27,28]. I consider class-structure at the patch level, in which patches
185 may be in different demographic states, but also class-structure at the individual level, in
186 which individuals within each patch may be in different states. I build upon recent work that
187 has sought to understand the evolution of social traits when patches vary in quality (e.g.
188 [15,17,31]), as well as when individuals within each patch vary in quality (e.g. [17,31]).

189

190 Fitness

191

192 To derive a fitness function in subdivided populations, I follow the formulation given in
193 Rodrigues and Gardner [17], which considers both variation in the demographic state of
194 patches across the population (i.e. between-patch heterogeneity) but also variation in
195 individual quality within each patch (i.e. within-patch heterogeneity). I partition fitness into
196 different additive components along three main dimensions. First, I consider reproductive
197 success according to whether this is achieved through a philopatric or dispersed component,
198 which I denote by the superscripts ‘ ϕ ’ and ‘ δ ’, respectively. Second, I consider class-structure

199 at the patch level, in which patches are classified according to their state $\sigma \in \Omega$, where the set
 200 Ω comprises all the possible patch demographic states, and where the state of patches can be
 201 defined according to different environmental or demographic variables, such as resource-
 202 availability [15,17], patch size [26], or age composition [35]. Finally, I consider class-
 203 structure within each patch, in which each individual is classified according to its quality $\rho \in$
 204 Ω_σ , where the set Ω_σ comprises all the possible individual qualities present in type- σ patches,
 205 and where an individual's quality can be defined, for instance, according to its future
 206 fecundity [31], maternal rank [17], or age [35]. I can then write the fitness of a focal class- ρ
 207 individual in a type- σ patch as

208

$$209 \quad W_{\rho,\sigma} = \frac{1}{v_{\rho,\sigma}} \left(\sum_{\sigma' \in \Omega} p_{\sigma \rightarrow \sigma'} \sum_{\rho' \in \Omega_{\sigma'}} w_{\rho \rightarrow \rho', \sigma \rightarrow \sigma'}^\phi v_{\rho', \sigma'} \right. \\ \left. + \sum_{\tau \in \Omega} p_\tau \sum_{\tau' \in \Omega} p_{\tau \rightarrow \tau'} \sum_{\rho' \in \Omega_{\tau'}} w_{\rho \rightarrow \rho', \sigma \rightarrow \tau \rightarrow \tau'}^\delta v_{\rho', \tau'} \right) \quad (6)$$

210

211 where: $p_{\sigma \rightarrow \sigma'}$ (or $p_{\tau \rightarrow \tau'}$) is the probability that a type- σ (or type- τ) patch becomes a type- σ' (or
 212 type- τ') patch; $w_{\rho \rightarrow \rho', \sigma \rightarrow \sigma'}^\phi$ is the philopatric component of the reproductive success of a focal
 213 class- ρ individual when it produces class- ρ' individuals associated with the demographic
 214 transition $\sigma \rightarrow \sigma'$; $v_{\rho', \sigma'}$ (or $v_{\rho', \tau'}$) is the reproductive value of a class- ρ' individual in a type- σ' (or
 215 type- τ') patch; p_τ is the frequency of type- τ patches in the population; and $w_{\rho \rightarrow \rho', \sigma \rightarrow \tau \rightarrow \tau'}^\delta$ is the
 216 reproductive success of a class- ρ individual in a type- σ patch when it produces class- ρ'
 217 individuals that disperse and arrive at type- τ patches that become type- τ' patches.

218

219 I assume that social interactions unfold within the local population, and therefore, I write
 220 down the fitness of a random individual in the local patch as

221

222 $w_\sigma = \sum_{\rho \in \Omega_\sigma} u_{\rho,\sigma} v_{\rho,\sigma} W_{\rho,\sigma},$ (7)

223

224 where the frequency of type- ρ individuals in a type- σ patch is given by $u_{\rho,\sigma} =$

225 $p_\sigma n_{\rho,\sigma} / \sum_{\tau \in \Omega} n_\tau p_\tau,$ in which $n_{\rho,\sigma}$ is the number of type- ρ individuals in a type- σ patch, and n_τ

226 is the total number of individuals in a type- τ patch. The fitness of a random individual in the

227 population is then given by

228

229 $w = \sum_{\sigma \in \Omega} w_\sigma.$ (8)

230

231 Behaviour conditional on individual quality

232

233 Let us now consider the evolution of a social trait whose average level of expression is given

234 by z . My aim is to obtain the fitness effects of a slight increase in trait value. I focus on a focal

235 recipient whose reproductive success is mediated by her own phenotype $x_{\rho,\sigma}$, encoded by the

236 breeding value $g_{\rho,\sigma}$, and by the phenotype of its neighbours $Y_{\alpha,\sigma}$, encoded by the breeding

237 value $G_{\alpha,\sigma}$ [6,17]. Actors are class- α individuals, while primary recipients (c.f. [31]) are class

238 $\rho \in \Theta$ individuals, in which Θ is the set that comprises all the classes that contain the primary

239 recipients of the behaviour, which may include the actor's class. To determine the effect of a

240 slight increase in breeding value on the fitness of the recipients in a type- σ patch, I take the

241 derivative of fitness with respect to the breeding value. This is given by

242

243
$$\frac{dw_\sigma}{dg_{\alpha,\sigma}} = \sum_{\rho \in \Omega_\sigma} u_\rho \left(\sum_{\sigma' \in \Omega} p_{\sigma \rightarrow \sigma'} \sum_{\rho' \in \Omega_{\sigma'}} \left(\frac{\partial w_{\rho \rightarrow \rho', \sigma \rightarrow \sigma'}^\phi}{\partial x_{\alpha,\sigma}} \frac{dx_{\alpha,\sigma}}{dg_{\alpha,\sigma}} \frac{dg_{\alpha,\sigma}}{dg_{\rho,\sigma}} + \right. \right.$$

$$\left. \frac{\partial w_{\rho \rightarrow \rho', \sigma \rightarrow \sigma'}^\phi}{\partial Y_{\alpha,\sigma}} \frac{dY_{\alpha,\sigma}}{dG_{\alpha,\sigma}} \frac{dG_{\alpha,\sigma}}{dg_{\rho,\sigma}} \right) v_{\rho',\sigma'} + \sum_{\tau \in \Omega} p_\tau \sum_{\tau' \in \Omega} p_{\tau \rightarrow \tau'} \sum_{\rho' \in \Omega_{\tau'}} \left(\frac{\partial w_{\rho \rightarrow \rho', \sigma \rightarrow \tau'}^\delta}{\partial x_{\alpha,\sigma}} \frac{dx_{\alpha,\sigma}}{dg_{\alpha,\sigma}} \frac{dg_{\alpha,\sigma}}{dg_{\rho,\sigma}} + \right.$$

$$\left. \frac{\partial w_{\rho \rightarrow \rho', \sigma \rightarrow \tau'}^\delta}{\partial Y_{\alpha,\sigma}} \frac{dY_{\alpha,\sigma}}{dG_{\alpha,\sigma}} \frac{dG_{\alpha,\sigma}}{dg_{\rho,\sigma}} \right) v_{\rho',\tau'} + \sum_{\sigma' \in \Omega} \frac{\partial p_{\sigma \rightarrow \sigma'}}{\partial Y_{\alpha,\sigma}} \frac{dY_{\alpha,\sigma}}{dG_{\alpha,\sigma}} \frac{dG_{\alpha,\sigma}}{dg_{\rho,\sigma}} \sum_{\rho' \in \Omega_{\sigma'}} W_{\rho \rightarrow \rho', \sigma \rightarrow \sigma'}^\phi v_{\rho',\sigma'} \Big)$$
 (9)

244

245 The partial derivatives are evaluated at the population's average trait value $z_{\rho,\sigma}$, and they
246 represent the effect of the phenotype on the reproductive success of each recipient ($\partial w/\partial x$),
247 and on the demographic state of the patch ($\partial p/\partial x$). The derivatives of the actor's breeding
248 value with respect to the recipient's breeding value, $dg_{\alpha,\sigma}/dg_{\rho,\sigma}$ and $dg_{\alpha,\sigma}/dG_{\rho,\sigma}$, represent the
249 coefficients of consanguinity between the actor and the recipients, denoted by $g_{\alpha\rho,\sigma}$, and the
250 coefficient of consanguinity between the actor and herself, denoted by $g_{\alpha^*,\sigma}$ [38]. The
251 derivatives of the phenotypes with respect to the breeding values represent the mapping
252 between the genotype and the phenotype, which we can set to one (i.e. $dx/dg = dY/dG = 1$).

253

254 The coefficients of relatedness between the actor α and the recipient ρ in a type- σ patch,
255 denoted by $R_{\alpha\rho,\sigma}$, is given by the ratio of the coefficient of consanguinity between the actor α
256 and recipient ρ ($g_{\alpha\rho,\sigma}$) and the coefficient of consanguinity between the actor α and herself
257 ($g_{\alpha^*,\sigma}$; [17,38]). Thus, $R_{\alpha\rho,\sigma} = g_{\alpha\rho,\sigma}/g_{\alpha^*,\sigma}$.

258

259 Behaviour conditional on patch quality

260

261 So far, we considered cases in which individuals express a trait according to their quality. In
262 some cases, however, information regarding individual quality may be unavailable, and
263 consequently, individuals may have to resort to a strategy conditional on the quality of the
264 local environment. Under this scenario, all individuals in the local patch are simultaneously
265 actors and recipients, irrespective of their individual quality (e.g. [15,31]). As a result, the
266 class of actors and recipients comprises all the classes present in the focal type- σ patch ($\Theta =$
267 Ω_σ). The fitness effect of the trait is then given by

268

$$269 \quad \frac{dw_\sigma}{dg_\sigma} = \sum_{\alpha \in \Omega_\sigma} \left. \frac{dw_\sigma}{dg_{\alpha,\sigma}} \right|_{x_{\alpha,\sigma}=Y_{\alpha,\sigma}=z_{\alpha,\sigma}, z_{\alpha,\sigma}=z_\sigma}, \quad (10)$$

270

271 where all the partial derivatives are evaluated at the average level of the behaviour in type- σ
272 patches (i.e. $x_{\alpha,\sigma} = Y_{\alpha,\sigma} = z_{\alpha,\sigma}$ and $z_{\alpha,\sigma} = z_\sigma$).

273

274 Unconditional behaviour

275

276 In some cases, information regarding the state of the local patch may be unavailable to
277 individuals, in which case social actors must express the behaviour unconditionally. In such
278 cases, the behaviour must be considered across all patch types. The fitness effect of an
279 unconditionally expressed behaviour is then given by

280

$$281 \quad \frac{dw}{dg} = \sum_{\sigma \in \Omega} \left. \frac{dw_\sigma}{dg_\sigma} \right|_{x_{\alpha,\sigma}=Y_{\alpha,\sigma}=z_\sigma, z_\sigma=z}, \quad (11)$$

282

283 where the partial derivatives are all evaluated at the population's average trait value (i.e. $x_{\alpha,\sigma} =$
284 $Y_{\alpha,\sigma} = z_\sigma$ and $z_\sigma = z$).

285

286 **Demography and continuous-time models**

287

288 As shown in equation (9), kin selection depends on three key variables: first, the frequency of
289 individuals in each class (i.e. u), which can be calculated from the frequency of each patch
290 type (i.e. p ; e.g. 33-35); second, the reproductive value of each individual (i.e. v); and finally,
291 the relatedness among group mates (i.e. R). Here, I provide a formal link between the

292 mathematics of stochastic continuous-time processes [39-42] and the mathematics of kin
293 selection.

294

295 Continuous-times models: general principles

296

297 Let us consider a discrete state space and continuous time, in which the stochastic process is
298 defined by $\{X(t) \in \Omega: t \in [0, \infty)\}$, where $X(t)$ is a discrete random variable with values defined
299 by the set Ω , and where the argument t is continuous. The stochastic process $\{X(t) \in \Omega: t \in$
300 $[0, \infty)\}$ is denominated a continuous-time Markov chain (CTMC) and is characterised by the
301 Markovian property, such that its current state depends only on the previous state and not on
302 the states leading to the previous state. The random variable $X(t)$ is described by the
303 probability distribution $\{p_\sigma(t)\}_{\sigma \in \Omega}$, where $p_\sigma(t) = \text{Prob}\{X(t) = \sigma\}$. The transition
304 probabilities between state σ and state τ are defined by $p_{\sigma \rightarrow \tau}(t, t_0) =$
305 $\text{Prob}\{X(t) = \tau | X(t_0) = \sigma\}$, $t_0 < t$, and I assume that the transition probabilities are
306 homogenous such that

307

308 $p_{\sigma \rightarrow \tau}(t - t_0) = \text{Prob}\{X(t) = \tau | X(t_0) = \sigma\} = \text{Prob}\{X(t - t_0) = \tau | X(0) = \sigma\},$ (12)

309

310 and therefore the transition probabilities depend only on the length of the time interval. The
311 transition matrix is defined as $\mathbf{P}(t) = (p_{\sigma \rightarrow \tau}(t))_{\xi \times \xi}$, where ξ is the number of elements in
312 the set Ω , and where the transition probabilities obey to the following property:

313 $\sum_{\tau \in \Omega} p_{\sigma \rightarrow \tau}(t) = 1$. The sequence of demographic states over time can be represented by the

314 directed graph of the embedded Markov chain (see Figure 1 for an example). Let us define

315 $\text{Dist}(\sigma, \tau)$ as the number of edges required to go from the demographic state σ to the

316 demographic state τ given the shortest path between them. If we consider a Poisson process,
 317 for a sufficiently small time interval Δt we have

318

$$\begin{aligned}
 p_{\sigma \rightarrow \sigma'}(\Delta t) &= \text{Prob}\{X(t + \Delta t) = \sigma' | X(t) = \sigma\} \\
 319 \quad &= \begin{cases} p_{\sigma \rightarrow \tau} \Delta t + o(\Delta t), & \sigma' = \tau \in \Omega_{\sigma}^1, \tau \neq \sigma \\ 1 - \sum_{\tau \in \Omega, \tau \neq \sigma} p_{\sigma \rightarrow \tau} \Delta t + o(\Delta t) & \sigma' = \sigma \\ o(\Delta t) & \sigma' = \eta \in \Omega_{\sigma}^{>1} \end{cases}, \quad (13)
 \end{aligned}$$

320

321 in which $o(\Delta t)$ – “little oh Δt ” – approaches zero more rapidly than Δt when $\Delta t \rightarrow 0$, and
 322 where the set Ω_{σ}^1 comprises the states that satisfy the condition $\text{Dist}(\sigma, \tau) = 1$, whereas the set
 323 $\Omega_{\sigma}^{>1}$ comprises the states that satisfy the condition $\text{Dist}(\sigma, \eta) > 1$.

324

325 Frequency of demographic states

326

327 As shown above, the fitness effects depend on the frequency of individuals in each class,
 328 which can be calculated from the frequency of patches in each state. Our CTMC is defined by
 329 equations (13), and therefore the probability that the Markov chain is in the state σ at time
 330 $t + \Delta t$ is given by

331

$$332 \quad p_{\sigma}(t + \Delta t) = \sum_{\tau \in \Omega, \tau \neq \sigma} p_{\tau}(t) p_{\tau \rightarrow \sigma} \Delta t + p_{\sigma}(t) \left(1 - \sum_{\tau \in \Omega, \tau \neq \sigma} p_{\sigma \rightarrow \tau} \Delta t\right) + \varepsilon o(\Delta t), \quad (14)$$

333

334 where $\varepsilon o(\Delta t)$ comprises all the terms in $o(\Delta t)$. Subtracting $p_{\sigma}(t)$ from both sides of the
 335 equation, dividing by Δt , and taking the limit $\Delta t \rightarrow 0$, we get the forward Kolmogorov
 336 differential equations, which are given by

337

$$338 \quad \frac{dp_{\sigma}}{dt} = \sum_{\tau \in \Omega, \tau \neq \sigma} p_{\tau \rightarrow \sigma} p_{\tau} - \sum_{\tau \in \Omega, \tau \neq \sigma} p_{\sigma \rightarrow \tau} p_{\sigma} \quad , \sigma \in \Omega, \quad (15)$$

339

340 where the positive terms represent cases in which a patch in the demographic state $\tau \neq \sigma$
341 becomes a type- σ patch, while the negative terms represent those cases in which a patch in the
342 demographic state σ becomes a type- τ ($\neq \sigma$) patch, with $\tau \in \Omega$. Over time, the system of
343 differential equations, for most cases, tends to an equilibrium state, which represents the
344 asymptotic probabilities that the patch is in any of the given demographic states. We can
345 represent this system of equations in a matrix form, by defining the generator matrix $\mathbf{Q} =$
346 $\lim_{\Delta t \rightarrow 0} \frac{\mathbf{P}(\Delta t) - \mathbf{I}}{\Delta t}$, where \mathbf{I} is the identity matrix. The forward Kolmogorov differential equations are
347 then given by $\frac{d\mathbf{p}}{dt} = \mathbf{Q}\mathbf{p}$, where $\mathbf{p}(t) = (p_\sigma(t))_{\sigma \in \Omega}$ is the vector of probabilities. The vector of
348 probabilities provides a direct link to the infinite island model. The state of each island is
349 described by the continuous-time Markov chain $\{X(t) \in \Omega: t \in [0, \infty)\}$ defined above, and
350 therefore, the probabilities \mathbf{p} also give the frequency of each patch type in the population. The
351 frequency of each patch type at equilibrium is given by the solution of the forward
352 Kolmogorov differential equations. Note that the forward Kolmogorov differential equations
353 in (15) recover the equations used in previous studies (i.e. Alizon and Taylor [33]; Wild et al.
354 [34]; and Rodrigues [35]).

355

356 Reproductive value

357

358 Second, the marginal fitness effects depend on the reproductive value of actors and recipients.
359 Reproductive value measures the value of an individual according to her capacity to send
360 copies of her genes to the gene pool of future generations [36,43,44]. Let us consider a very
361 short time interval Δt and how the reproductive value of class- α individuals in a type- σ patch
362 changes over this time interval. From the fitness function (equation 6) and the transition
363 probabilities of the CTMC (equations 13), the change in reproductive value is given by

364

$$\begin{aligned}
 365 \quad v_{\alpha,\sigma}(t + \Delta t) &= \sum_{\sigma' \in \Omega} p_{\sigma \rightarrow \sigma'}(\Delta t) \sum_{\alpha' \in \Omega_{\sigma'}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \sigma'}^{\phi} v_{\alpha', \sigma'}(t) \\
 &+ \sum_{\tau \in \Omega} p_{\tau} \sum_{\tau' \in \Omega} p_{\tau \rightarrow \tau'}(\Delta t) \sum_{\alpha' \in \Omega_{\tau}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \tau \rightarrow \tau'}^{\delta} v_{\alpha', \tau'}(t)
 \end{aligned} \tag{16}$$

366

367 Expanding the right-hand side of this equation, and given that $w_{\alpha \rightarrow \alpha, \sigma \rightarrow \sigma}^{\phi} = 1$, we obtain

368

$$\begin{aligned}
 369 \quad v_{\alpha,\sigma}(t + \Delta t) &= \sum_{\sigma' \in \Omega, \sigma' \neq \sigma} p_{\sigma \rightarrow \sigma'} \Delta t \left(\sum_{\alpha' \in \Omega_{\sigma'}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \sigma'}^{\phi} v_{\alpha', \sigma'}(t) - v_{\alpha,\sigma}(t) \right) + v_{\alpha,\sigma}(t) + \\
 370 \quad &\sum_{\tau \in \Omega} p_{\tau} \sum_{\tau' \in \Omega} p_{\tau \rightarrow \tau'}(\Delta t) \sum_{\alpha' \in \Omega_{\tau}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \tau \rightarrow \tau'}^{\delta} v_{\alpha', \tau'}(t) + \varepsilon o(\Delta t).
 \end{aligned} \tag{17}$$

371

372 Subtracting $v_{\alpha,\sigma}(t)$ from both sides of the equation, dividing by Δt , and taking the limit $\Delta t \rightarrow$

373 0, we get the forward Kolmogorov differential equation for reproductive value, which is given

374 by

375

$$\begin{aligned}
 376 \quad \frac{dv_{\alpha,\sigma}}{dt} &= \sum_{\sigma' \in \Omega, \sigma' \neq \sigma} p_{\sigma \rightarrow \sigma'} \left(\sum_{\alpha' \in \Omega_{\sigma'}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \sigma'}^{\phi} v_{\alpha', \sigma'} - v_{\alpha,\sigma} \right) \\
 &+ \sum_{\tau \in \Omega} p_{\tau} \sum_{\tau' \in \Omega} p_{\tau \rightarrow \tau'} \sum_{\alpha' \in \Omega_{\tau}} w_{\alpha \rightarrow \alpha', \sigma \rightarrow \tau \rightarrow \tau'}^{\delta} v_{\alpha', \tau'}(t) \quad , \sigma \in \Omega, \alpha \in \Omega_{\sigma}
 \end{aligned} \tag{18}$$

377

378 The system of differential equations describes the change in the reproductive value of each

379 individual over time. At equilibrium, reproductive value will remain constant. We can then set

380 the derivatives to zero, and solve the system of equations to obtain the reproductive value of

381 each individual in a normal population (i.e. the mutant allele is neutral; [10,44]).

382

383 Relatedness

384

385 Finally, the marginal fitness effects depend on the coefficient of relatedness among social

386 partners, which gives the evolutionary measure of value an actor employs to evaluate social

387 partners according to her genetic interests [1,2]. We can derive the coefficients of relatedness
 388 from the coefficients of consanguinity between social partners [16,17,38], and from the
 389 backwards probabilities associated with the demographic transitions in patch state (e.g.
 390 [15,17]). Specifically, the probability that given the demographic state σ the patch was in the
 391 demographic state l some Δt time units ago is given by

$$\begin{aligned}
 p_{\sigma'|\sigma}(\Delta t) &= \text{Prob}\{X(t - \Delta t) = \sigma' | X(t) = \sigma\} \\
 393 \quad &= \begin{cases} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} \Delta t + o(\Delta t), & \sigma' = \tau \in \Omega_{\sigma}^1, \tau \neq \sigma \\ 1 - \sum_{\tau \in \Omega, \tau \neq \sigma} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} \Delta t + o(\Delta t) & \sigma' = \sigma \\ o(\Delta t) & \sigma' = \eta \in \Omega_{\sigma}^{>1} \end{cases} \quad (19)
 \end{aligned}$$

394
 395 Thus, the change in the coefficient of consanguinity between class- ρ and class- η
 396 individuals in a type- σ patch over a short time interval Δt is given by

$$\begin{aligned}
 397 \quad g_{\rho\eta,\sigma}(t + \Delta t) &= \sum_{\tau \in \Omega, \tau \neq \sigma} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} \Delta t g_{\rho\eta:\sigma,\tau \rightarrow \sigma}(t) \\
 398 \quad &+ \left(1 - \sum_{\tau \in \Omega, \tau \neq \sigma} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} \Delta t\right) g_{\rho\eta:\sigma}(t) + \varepsilon o(\Delta t) \quad (20)
 \end{aligned}$$

399
 400 where $g_{\rho\eta:\sigma,\tau \rightarrow \sigma}(t)$ is the coefficient of consanguinity in a type- σ patch after the patch has
 401 transitioned from state τ to state σ . Note that if a type- σ patch has kept its demographic state
 402 over the Δt time period, then the coefficient of consanguinity is simply $g_{\rho\eta:\sigma}$. Subtracting
 403 $g_{\rho\eta:\sigma}(t)$ from both sides of the equation, dividing by Δt , and taking the limit $\Delta t \rightarrow 0$, we get
 404 the differential equations for the coefficient of consanguinities, which is given by

$$406 \quad \frac{dg_{\rho\eta,\sigma}}{dt} = \sum_{\tau \in \Omega, \tau \neq \sigma} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} g_{\rho\eta:\sigma,\tau \rightarrow \sigma} - \sum_{\tau \in \Omega, \tau \neq \sigma} \frac{p_{\tau \rightarrow \sigma} p_{\tau}}{p_{\sigma}} g_{\rho\eta,\sigma} \quad \rho, \eta \in \Omega_{\sigma}, \quad (21)$$

407

408 The system of equations can be solved to find the coefficients of consanguinity, which at
409 equilibrium remains constant over time.

410

411 **Hamilton's rule and inclusive fitness**

412

413 Here, I derive a general formulation of Hamilton's rule for a range of social traits. Hamilton's
414 rule enables an inclusive fitness interpretation of the behaviour, and therefore it establishes a
415 formal link between the neighbour-modulated approach and Hamilton's inclusive fitness
416 theory [1,6,18]. I first analyse a helping behaviour that influences the fecundity of the actor
417 and recipients, followed by the analysis of a helping behaviour that influences the survival of
418 the actor and recipients. In Appendix A, I show how to extend the framework to analyse the
419 evolution of a dispersal trait. For each trait, I consider behaviours that are expressed
420 conditionally on the quality of individuals, but also traits that are expressed conditionally on
421 patch quality as well as traits that are expressed unconditionally.

422

423 General life-cycle

424

425 I consider a general model in which the fecundity rate of class- ρ mothers in type- σ patches is
426 given by $f_{\rho,\sigma}$. Offspring remain in the local patch with probability $1 - d_{\rho,\sigma}$, and disperse to a
427 random patch in the population with probability $d_{\rho,\sigma}$. Offspring occupy empty breeding sites
428 in type- σ patches at a rate o_{σ} , and offspring who fail to acquire a breeding site die. Resident
429 mothers die at a rate $m_{\rho,\sigma}$. In addition, I consider that resident mothers may undergo changes
430 in their reproductive, physiological, or social state. Such cases are represented by
431 demographic rates of the form $a_{\rho \rightarrow \eta}$, which specify the rate at which class- ρ mothers become
432 class- η mothers.

433

434 *Demographic rates* -- Given these general model assumptions, let us consider key
435 demographic rates. First, the demographic state of a focal type- σ patch may change whenever
436 a mother dies. This is given by

437

$$438 \quad p_{\sigma \rightarrow \sigma; \rho; \dagger} = n_{\rho, \sigma} m_{\rho^{\circ}, \sigma}, \quad (22)$$

439

440 where: $m_{\rho^{\circ}, \sigma}$ is the average mortality rate of class- ρ mothers in type- σ patches; and $\sigma; \rho; \dagger$ is the
441 demographic state of a type- σ patch after a class- ρ mother has died. Second, the demographic
442 state of a focal patch may change whenever an offspring occupies an empty breeding site.

443 This is given by

444

$$445 \quad p_{\sigma \rightarrow \sigma; \emptyset} = \sum_{\rho \in \Omega_{\sigma}} n_{\rho, \sigma} f_{\rho^{\circ}, \sigma} (1 - d_{\rho^{\circ}, \sigma}) + D, \quad (23)$$

446

447

448 where: $f_{\rho^{\circ}, \sigma}$ is the average fecundity of class- ρ individuals in the focal type- σ patch; $d_{\rho^{\circ}, \sigma}$ is
449 the average level of dispersal of class- ρ individuals in the focal type- σ patch; $D =$
450 $\sum_{\tau \in \Omega} (\sum_{\eta \in \Omega_{\tau}} n_{\eta, \tau} f_{\eta, \tau} d_{\eta, \tau})$ is the total number of immigrants that arrive at the focal patch, in
451 which $d_{\eta, \tau}$ is the population's average dispersal rate of class- η individuals in type- τ patches;
452 and $\sigma; \emptyset$ is the demographic state of a type- σ patch after a new offspring has occupied a vacant
453 breeding site.

454

455 *Reproductive success* -- We now need to determine the reproductive success of a focal mother
456 associated with each of the demographic rates. First, let us consider the reproductive success
457 of a resident mother when a class- ρ mother dies. If the mother is not a class- ρ mother, then

458 her reproductive success is necessarily one. Otherwise, the reproductive success of a focal
 459 class- ρ mother is given by

460

$$461 \quad w_{\rho \rightarrow \rho, \sigma \rightarrow \sigma: \rho: \dagger}^{\phi} = 1 - \frac{m_{\rho^*, \sigma}}{n_{\rho, \sigma} m_{\rho^*, \sigma}}, \quad (24)$$

462

463 where $w_{\rho \rightarrow \rho, \sigma \rightarrow \sigma: \rho: \dagger}^{\phi}$ is the reproductive success of the class- ρ mother when a class- ρ mother
 464 dies in the focal type- σ patch, and $m_{\rho^*, \sigma}$ is the mortality rate of the focal class- ρ mother in the
 465 focal type- σ patch. Let us now consider the reproductive success of a resident mother when an
 466 offspring occupies a vacant breeding site. This has two additive components. First, the focal
 467 mother generates reproductive success via her own survival, which is necessarily one, i.e.

468 $w_{\rho \rightarrow \rho, \sigma \rightarrow \sigma: \emptyset}^{\phi} = 1$. Second, the focal mother generates reproductive success if the offspring that

469 occupies the empty breeding site in the local patch is her own. This is given by

470

$$471 \quad w_{\rho \rightarrow \emptyset, \sigma \rightarrow \sigma: \emptyset}^{\phi} = \frac{f_{\rho^*, \sigma} (1 - d_{\rho^*, \sigma})}{\sum_{\eta \in \Omega_{\sigma}} n_{\eta, \sigma} f_{\eta^*, \sigma} (1 - d_{\eta^*, \sigma}) + D}, \quad (25)$$

472

473 where: $f_{\rho^*, \sigma}$ is the fecundity of the focal class- ρ recipient in the focal type- σ patch; $d_{\rho^*, \sigma}$ is the

474 probability of dispersal of the focal mother's offspring; and $w_{\rho \rightarrow \emptyset, \sigma \rightarrow \sigma: \emptyset}^{\phi}$ is the reproductive

475 success of a focal mother via the production of philopatric offspring. Third, the focal mother

476 generates reproductive success if her offspring occupy an empty breeding site in a foreign

477 patch. This is given by

478

$$479 \quad w_{\rho \rightarrow \emptyset, \sigma \rightarrow \tau \rightarrow \tau': \emptyset}^{\delta} = \frac{f_{\rho^*, \sigma} d_{\rho^*, \sigma}}{\sum_{\eta \in \Omega_{\tau}} n_{\eta, \tau} f_{\eta^*, \tau} (1 - d_{\eta^*, \tau}) + D}. \quad (26)$$

480

481 Behaviour conditional on individual quality

482

483 *Helping: fecundity effects* -- Let us consider a general helping trait that reduces the fecundity
 484 rate of the actor but raises the fecundity of the actor's social partners, and that is conditionally
 485 expressed on the quality of the actor and recipients. I assume that actors are members of the
 486 class- α , and recipients are members of the class $\beta \in \Theta$, where Θ is the set of the classes that
 487 include the primary recipients of the behaviour (i.e. those that receive the immediate benefits
 488 of the behaviour). The fecundity of a focal class- ρ recipient is given by $f_{\rho,\sigma} = f_{\rho,\sigma}(x_{\rho,\sigma}, Y_{\alpha,\sigma})$,
 489 while the average fecundity of the class- ρ recipients is given by $f_{\rho,\sigma} = f_{\rho,\sigma}(Y_{\rho,\sigma}, Y_{\alpha,\sigma})$. All the
 490 other traits are set to the population's average values. Because the behaviour influences the
 491 fecundity of each recipient, I focus on the demographic transitions and class-specific
 492 reproductive successes that depend on fecundity only, as outlined above. Replacing the
 493 expressions (22-26) in equation (9), we get

494

$$495 \frac{dW_{\alpha,\sigma}}{dg_{\alpha,\sigma}} = u_{\alpha,\sigma} \left(\frac{\partial f_{\alpha,\sigma}(x_{\alpha,\sigma}, Y_{\alpha,\sigma})}{\partial x_{\alpha,\sigma}} V_{\alpha,\sigma} + \sum_{\rho \in \Theta} \frac{n_{\rho,\sigma}}{n_{\alpha,\sigma}} \frac{\partial f_{\rho,\sigma}(x_{\rho,\sigma}, Y_{\alpha,\sigma})}{\partial Y_{\alpha,\sigma}} V_{\rho,\sigma} R_{\alpha\rho,\sigma} + \right. \\ \left. \sum_{\eta \in \Theta} \frac{n_{\eta,\sigma}}{n_{\alpha,\sigma}} \frac{\partial f_{\eta,\sigma}(Y_{\eta,\sigma}, Y_{\alpha,\sigma})}{\partial Y_{\alpha,\sigma}} (1 - d_{\eta,\sigma}) o_{\sigma} \sum_{\rho \in \Omega_{\sigma}} n_{\rho,\sigma} R_{\alpha\rho,\sigma} (v_{\rho,\sigma;\emptyset} - v_{\rho,\sigma}) \right) \quad (27)$$

496

497 where $V_{\alpha,\sigma}$ (or $V_{\rho,\sigma}$) is the reproductive value of an actor's (or recipient's) offspring, and $v_{\rho,\sigma;\emptyset}$
 498 is the reproductive value of a type- ρ individual when the focal type- σ patch accommodates a
 499 new breeder. The coefficient of relatedness among members of the same class $R_{\alpha\alpha,\sigma}$ is the
 500 'whole-group' coefficient of relatedness, which is given by $R_{\alpha\alpha,\sigma} = 1/n_{\alpha,\sigma} + ((n_{\alpha,\sigma} - 1)/n_{\alpha,\sigma})r_{\alpha\alpha,\sigma}$,
 501 where $r_{\alpha\alpha,\sigma}$ is the 'others-only' coefficient of relatedness [45]. The reproductive value of an
 502 offspring is given by

503

$$504 \quad V_{\alpha,\sigma} = \frac{1}{f_{\alpha,\sigma}} \left(p_{\sigma \rightarrow \sigma:\emptyset} W_{\alpha \rightarrow \emptyset, \sigma \rightarrow \sigma:\emptyset}^{\phi} v_{\emptyset, \sigma:\emptyset} + \sum_{\tau \in \Omega} p_{\tau} p_{\tau \rightarrow \tau:\emptyset} W_{\alpha \rightarrow \emptyset, \sigma \rightarrow \tau:\emptyset}^{\delta} v_{\emptyset, \tau:\emptyset} \right). \quad (28)$$

505

506 where $v_{\emptyset, \sigma:\emptyset}$ (or $v_{\emptyset, \tau:\emptyset}$) is the reproductive value of an offspring that has taken up a breeding
 507 spot in a type- σ (or type- τ) patch. Note that $\partial f_{\eta, \sigma}(x_{\eta, \sigma}, Y_{\alpha, \sigma}) / \partial Y_{\alpha, \sigma}$ is the effect of the class- α
 508 actors on the fecundity of a focal class- η recipient. Thus, the effect of a single class- α focal
 509 actor on class- η recipients is given by $(n_{\eta, \sigma} / n_{\alpha, \sigma}) (\partial f_{\eta, \sigma} / \partial Y_{\alpha, \sigma})$, which gives the inclusive fitness
 510 fecundity benefit $B_{\alpha \eta, \sigma}$ provided to the class- η recipients by a focal class- α actor. Moreover,
 511 $\partial f_{\alpha, \sigma}(x_{\alpha, \sigma}, Y_{\alpha, \sigma}) / \partial x_{\alpha, \sigma}$ is the effect of a class- α actor on her own fecundity, where the additive
 512 inverse gives the inclusive fitness fecundity cost $C_{\alpha, \sigma}$ to the actor. Replacing these variables in
 513 equation (27), we get Hamilton's rule, which is given by

514

$$515 \quad \begin{aligned} & -C_{\alpha, \sigma} V_{\alpha, \sigma} + \sum_{\rho \in \Theta} B_{\alpha \rho, \sigma} V_{\rho, \sigma} R_{\alpha \rho, \sigma} \\ & + C_{\alpha, \sigma} (1 - d_{\alpha, \sigma}) o_{\sigma} \sum_{\rho \in \Omega_{\sigma}} n_{\rho, \sigma} R_{\alpha \rho, \sigma} (v_{\rho, \sigma} - v_{\rho, \sigma:\emptyset}) \\ & + \sum_{\eta \in \Theta} B_{\alpha \eta, \sigma} (1 - d_{\eta, \sigma}) o_{\sigma} \sum_{\rho \in \Omega_{\sigma}} n_{\rho, \sigma} R_{\alpha \rho, \sigma} (v_{\rho, \sigma:\emptyset} - v_{\rho, \sigma}) > 0 \end{aligned} \quad (29)$$

516

517 The left-hand side of this inequality immediately yields an inclusive fitness interpretation of
 518 the behaviour. A focal actor has $C_{\alpha, \sigma}$ fewer offspring, whose reproductive value is $V_{\alpha, \sigma}$. A
 519 class- ρ recipient enjoys a benefit that enables her to produce $B_{\alpha \rho, \sigma}$ extra offspring, with each
 520 additional offspring leading to an increment $V_{\rho, \sigma}$ in reproductive value, that must be
 521 depreciated by the relatedness $R_{\alpha \rho, \sigma}$ between actor and recipient. Of the additional $B_{\alpha \eta, \sigma}$
 522 offspring, a fraction $1 - d_{\eta, \sigma}$ remain in the local patch, which results in an empty site being
 523 occupied with probability o_{σ} . The occupation of the vacant breeding site changes the
 524 reproductive value of all the $n_{\rho, \sigma}$ class- ρ resident adults in the patch, who see their
 525 reproductive value change from $v_{\rho, \sigma}$ to $v_{\rho, \sigma:\emptyset}$. These changes in reproductive value must be
 526 depreciated by the relatedness $R_{\alpha \rho, \sigma}$ between the actor and the recipients. Finally, the actor

527 produces $C_{\alpha,\sigma}$ fewer offspring, who with probability $1 - d_{\alpha,\sigma}$ would have stayed in the local
 528 patch, and with probability o_σ would have occupied a vacant breeding spot. This influences
 529 the demographic environment of the actor's social partners, including herself, who see their
 530 reproductive value change from $v_{\rho,\sigma;\emptyset}$ to $v_{\rho,\sigma}$. These changes in reproductive value must be
 531 depreciated by the relatedness $R_{\alpha\rho,\sigma}$ between the actor and recipients.

532

533 *Helping: survival effects* -- Let us now consider a general helping trait that reduces the
 534 survival of the actor, denoted by $s_{\alpha,\sigma}$, to improve the survival of the recipient(s), denoted by
 535 $s_{\rho,\sigma}$, in which $s_{\alpha,\sigma} = -m_{\alpha,\sigma}$ and $s_{\rho,\sigma} = -m_{\rho,\sigma}$. The survival of a focal class- ρ recipient is given by
 536 $s_{\rho,\sigma} = s_{\rho,\sigma}(x_{\rho,\sigma}, Y_{\alpha,\sigma})$, whilst the average fecundity of class- ρ recipients is given by $s_{\rho\circ,\sigma} =$
 537 $s_{\rho,\sigma}(Y_{\rho,\sigma}, Y_{\alpha,\sigma})$. All the other traits are set to the population's average values. As above, we only
 538 need to consider the demographic rates and reproductive successes that directly depend on the
 539 survival of individuals. If we plug in the expressions (22-26) in equation (9), we get the
 540 following equation for the marginal fitness effects of the behaviour

541

$$\begin{aligned}
 \frac{dW_{\alpha,\sigma}}{dg_{\alpha,\sigma}} = & u_{\alpha,\sigma} \left(\frac{\partial s_{\alpha,\sigma}(x_{\alpha,\sigma}, Y_{\alpha,\sigma})}{\partial x_{\alpha,\sigma}} v_{\alpha,\sigma} + \sum_{\rho \in \Theta} \frac{n_{\rho,\sigma}}{n_{\alpha,\sigma}} \frac{\partial s_{\rho,\sigma}(x_{\rho,\sigma}, Y_{\alpha,\sigma})}{\partial Y_{\alpha,\sigma}} R_{\alpha\rho,\sigma} v_{\rho,\sigma} + \frac{s_{\alpha,\sigma}(x_{\alpha,\sigma}, Y_{\alpha,\sigma})}{\partial x_{\alpha,\sigma}} \right. \\
 & (n_{\alpha,\sigma} - 1) r_{\alpha\alpha,\sigma} (v_{\alpha,\sigma} - v_{\alpha,\sigma;\alpha;\dagger}) + \frac{s_{\alpha,\sigma}(x_{\alpha,\sigma}, Y_{\alpha,\sigma})}{\partial x_{\alpha,\sigma}} \sum_{\rho \neq \alpha, \rho \in \Omega_\sigma} n_{\rho,\sigma} R_{\rho\alpha,\sigma} (v_{\rho,\sigma} - v_{\rho,\sigma;\alpha;\dagger}) \\
 & \left. + \sum_{\eta \in \Theta} \frac{n_{\eta,\sigma}}{n_{\alpha,\sigma}} \frac{s_{\eta,\sigma}(Y_{\eta,\sigma}, Y_{\alpha,\sigma})}{\partial Y_{\alpha,\sigma}} (n_{\eta,\sigma} - 1) R_{\alpha\eta,\sigma} (v_{\eta,\sigma} - v_{\eta,\sigma;\eta;\dagger}) + \right. \\
 & \left. \sum_{\eta \in \Theta} \frac{n_{\eta,\sigma}}{n_{\alpha,\sigma}} \frac{s_{\eta,\sigma}(Y_{\eta,\sigma}, Y_{\alpha,\sigma})}{\partial Y_{\alpha,\sigma}} \sum_{\rho \neq \eta, \rho \in \Omega_\sigma} n_{\rho,\sigma} R_{\rho\eta,\sigma} (v_{\rho,\sigma} - v_{\rho,\sigma;\eta;\dagger}) \right) \quad (30)
 \end{aligned}$$

543

544 where $v_{\alpha,\sigma}$ is the reproductive value of the actor α , and $v_{\alpha,\sigma;\alpha;\dagger}$ is the reproductive value of a
 545 class- α individual after a class- α social partner has died. The same notation applies to the
 546 other variables.

547

548 Let us now consider the inclusive fitness survival costs and benefits. First, $\partial S_{\alpha,\sigma}(x_{\alpha,\sigma}, Y_{\alpha,\sigma})/\partial x_{\alpha,\sigma}$
 549 represents the effect of the class- α actor's behaviour on her own survival, and consequently
 550 the additive inverse gives the inclusive fitness survival cost $C_{\alpha,\sigma}$ paid by the actor. Second,
 551 $\partial S_{\rho,\sigma}(Y_{\rho,\sigma}, Y_{\alpha,\sigma})/\partial Y_{\alpha,\sigma}$ represents the effects of the class- α actors' behaviour on the survival of
 552 the focal class- ρ recipient, and consequently $(n_{\rho,\sigma}/n_{\alpha,\sigma})(\partial S_{\rho,\sigma}/\partial Y_{\alpha,\sigma})$ gives the inclusive fitness
 553 fecundity benefit $B_{\alpha\rho}$ provided by a class- α actor to class- ρ recipients. Replacing these
 554 variables in equation (30), we get Hamilton's rule, which can be written as

555

$$\begin{aligned}
 & -C_{\alpha,\sigma}v_{\alpha,\sigma} + \sum_{\rho \in \Theta} B_{\alpha\rho,\sigma}R_{\alpha\rho,\sigma}v_{\rho,\sigma} - C_{\alpha,\sigma}(n_{\alpha,\sigma} - 1)r_{\alpha\alpha,\sigma}(v_{\alpha,\sigma} - v_{\alpha,\sigma:\alpha:\dagger}) - \\
 556 & C_{\alpha,\sigma} \sum_{\rho \neq \alpha, \rho \in \Omega_\sigma} n_{\rho,\sigma}R_{\rho\alpha,\sigma}(v_{\rho,\sigma} - v_{\rho,\sigma:\alpha:\dagger}) + \sum_{\eta \in \Theta} B_{\alpha\eta,\sigma}(n_{\eta,\sigma} - 1)R_{\alpha\eta,\sigma}(v_{\eta,\sigma} - v_{\eta,\sigma:\eta:\dagger}). \quad (31) \\
 & + \sum_{\eta \in \Theta} B_{\alpha\eta,\sigma} \sum_{\rho \neq \eta, \rho \in \Omega_\sigma} n_{\rho,\sigma}R_{\alpha\rho,\sigma}(v_{\rho,\sigma} - v_{\rho,\sigma:\eta:\dagger}) > 0
 \end{aligned}$$

557

558 The left-hand side of this inequality immediately yields an inclusive fitness interpretation of
 559 the behaviour. An actor pays a survival cost $C_{\alpha,\sigma}$ and consequently loses all of her
 560 reproductive value $v_{\alpha,\sigma}$. A class- ρ recipient enjoys a benefit $B_{\alpha\rho,\sigma}$, which leads to an increment
 561 $v_{\rho,\sigma}$ in her future reproductive value, an increment that must be depreciated by the relatedness
 562 $R_{\alpha\rho,\sigma}$ between the actor and the recipient. The death of the actor has an impact on the other $n_{\alpha,\sigma}$
 563 $- 1$ class- α individuals, who see their reproductive value change from $v_{\alpha,\sigma}$ to $v_{\alpha,\sigma:\alpha:\dagger}$. These
 564 changes in reproductive value must be depreciated by the relatedness $r_{\alpha\alpha,\sigma}$ between the actor
 565 and her social and class partners. The death of the actor also impacts individuals in other
 566 classes, who see their reproductive value change from $v_{\rho,\sigma}$ to $v_{\rho,\sigma:\alpha:\dagger}$. Moreover, the
 567 additional survival of a class- η individual has an impact on the other $n_{\eta,\sigma} - 1$ class- η
 568 individuals, who see their reproductive value change from $v_{\eta,\sigma:\eta:\dagger}$ to $v_{\eta,\sigma}$. These changes in
 569 reproductive value must be depreciated by the relatedness $R_{\alpha\eta,\sigma}$ between the actor and the
 570 class- η recipients. Furthermore, the additional survival of class- η individuals has an impact on

571 individuals in other classes, who see their reproductive value change from $v_{\rho,\sigma;\eta;\dagger}$ to $v_{\rho,\sigma}$.

572 These changes in reproductive value must be depreciated by the relatedness $R_{\alpha,\rho}$ between the

573 class- α actor and the class- ρ recipients.

574

575 Behaviour conditional on patch quality

576

577 The marginal fitness effects when behaviour is conditional on the state of the patch, but not

578 on the state of individuals, follows immediately from equation (10). We now consider that all

579 individuals are actors and recipients, irrespective of their class. Thus, the set of primary

580 recipients of the behaviour is equal to the set of all classes present in the focal patch, i.e. $\Theta =$

581 Ω_σ . In addition, the partial derivatives are now evaluated at the average trait value for the

582 type- σ patch, which is given by z_σ . Thus, $x_{\alpha,\sigma} = Y_{\alpha,\sigma} = z_{\alpha,\sigma}$ and $z_{\alpha,\sigma} = z_\sigma$, as shown in equation

583 (10). The inclusive fitness interpretation is similar to the one given above for behaviours

584 expressed conditionally on the quality of the individuals.

585

586 Unconditional behaviour

587

588 Unconditional traits are averages of the marginal effects occurring at the different patch types.

589 Hamilton's rule follows immediately from equation (11). The marginal fitness effects must

590 now be added together and the partial derivatives are evaluated at the population-wise average

591 levels. Thus, $x_\sigma = Y_\sigma = z_\sigma$. The inclusive fitness interpretation is similar to the one given above

592 for behaviour expressed conditionally on the quality of the individuals.

593

594

595

596 **Applications of the general framework**

597

598 Here, I analyse several examples to illustrate how the framework developed above can be
599 employed in the study of kin selection problems. In particular, I focus on three studies that
600 use the continuous-time framework but apply the inclusive-fitness method to kin selection to
601 investigate the adaptive evolution of social traits (Alizon and Taylor [33]; Wild *et al.* [34];
602 and Rodrigues [35]). I show how the theory of Markovian processes is connected to these
603 studies and how the general forms of Hamilton's rule derived above can be readily applied to
604 solve and analyse kin selection problems (see Appendix B for more details).

605

606 As we saw above from the general analysis of kin selection in a demographic context, fitness
607 depends on three key variables: first, the frequency of patches in each of the demographic
608 states; second, the reproductive value of each individual; and finally, the kin selection
609 coefficients of relatedness between social partners. From the analysis of stochastic processes,
610 we saw that these three quantities are calculated from the Kolmogorov differential equations
611 defined by the system of equations (15), (18), and (21). These equations have precisely the
612 same form of the equations used by Alizon and Taylor [33], Wild *et al.* [34], and Rodrigues
613 [35] to calculate each of the key quantities. Thus, the Kolmogorov differential equations
614 (15,18,21) provide a formal justification for the heuristics used in each of these three studies,
615 and establish a link between their derivations.

616

617 Let us now consider Hamilton's rule. Three key studies have employed the continuous-time
618 framework to study the evolution of different social traits. Alizon and Taylor [33] studied a
619 helping trait as a function of group size. Wild *et al.* [34] studied the evolution of parasite
620 virulence. Finally, I studied the evolution of age-dependent social behaviour [35]. These

621 studies generated a form of Hamilton's rule to determine the selection acting on the traits of
622 interest. In the appendix B, I show that each of these Hamilton's rule are particular cases of
623 the general formulations of Hamilton's rule derived above. Thus, these general forms of
624 Hamilton's rule can be employed to analyse a wide range of evolutionary problems and unify
625 different studies.

626

627 **Extensions of the framework: Transmission of individual quality**

628

629 So far, in the general formulation of our problem, I assumed that there was no correlation
630 between parental and offspring quality. In some cases, however, parents may transmit their
631 quality to their offspring (e.g. [6]). Here, I show that these cases can be easily incorporated
632 within the general framework. In particular, I consider a scenario where individuals are either
633 of high- or low-quality and offspring can inherit parental quality to variable degrees of
634 fidelity.

635

636 I assume that each patch has at most two resident breeders, who can be either high- or low-
637 quality. High-quality (or low-quality) breeders give birth at a rate f_1 (or f_2) and die at a rate μ_1
638 (or μ_2). Offspring of high-quality (or low-quality) breeders become high-quality with
639 probability α (or $1 - \beta$), and low-quality with probability $1 - \alpha$ (or β). Offspring remain in the
640 natal patch with probability $1 - d_{ij}$, and disperse with probability d_{ij} , where i is the quality of
641 the mother and j if the quality of the mother's social partner. When $j = 0$, then the quality- i
642 mother is the sole resident breeder of the focal patch. Offspring settle in empty breeding spots
643 at a rate $o_1 = l/n$, where l is the number of empty breeding spots in the focal patch, and n is the
644 total capacity of the patch (i.e. $l = \{0,1,2\}$ and $n = 2$). I consider a helping trait that affects the
645 survival of actor and recipient, and therefore Hamilton's rule is given by inequality (31). In

646 appendix C, I explain the methodology in detail and I use numerical methods to solve the
647 model (e.g. [33-35]).

648

649 I find that in pure patches (i.e. patches occupied by breeders of the same quality), there is an
650 increase in the potential for helping as the transmission of quality increases (Figure 2A).

651 However, while in pure patches composed of low-quality individuals the potential for helping

652 always increases with inheritance of quality, in pure patches composed of high-quality

653 individuals the potential for helping decreases as inheritance of quality nears one (Figure 2A).

654 This is because when transmission of quality increases, the density of high-quality immigrants

655 also increases, which leads to a decrease in relatedness in patches composed of two high-

656 quality individuals, but an increase in relatedness in patches composed of two low-quality

657 individuals (Figure C1.A in Appendix C).

658

659 In mixed patches, there is a tendency for the potential for helping to decrease with the fidelity

660 of inheritance (Figure 2B). This is because as the fidelity of transmission increases, the

661 relatedness among social partners decreases (Figure C1.B). I also find that high-quality

662 breeders are always selected to invest less in helping than low-quality breeders. This occurs

663 because high-quality breeders have higher reproductive value than low-quality breeders

664 (Figure C1.D). Finally, I find that as the fidelity of transmission approach one, the potential

665 for helping of low-quality breeders increases with the fidelity of transmission (Figure 2B).

666 This is because when the transmission of quality nears one, the reproductive value of low-

667 quality breeders rapidly decreases (Figure C1.D).

668

669 We now turn the attention to the evolution of unconditional behaviour. I find that the potential

670 for helping is independent of the fidelity of transmission (Figure 2C). While the potential for

671 helping increases with the fidelity of transmission in pure patches, it decreases in mixed
672 patches, such that average potential for helping depends little on the fidelity of transmission.

673

674 **Discussion**

675

676 Here, I provided a framework connecting the formal theory of kin selection with the
677 fundamental theory of stochastic continuous-time models. First, I derived a general
678 expression of fitness in a demographic context. Second, I provided a link between the
679 fundamental theory of stochastic processes, demography, and key kin selection variables.
680 Finally, I derived Hamilton's rule for major social traits. Specifically, I employed the
681 neighbour-modulated approach to kin selection to provide a general account of inclusive
682 fitness in continuous-time models. I partitioned the neighbour-modulated fitness of a focal
683 recipient to account for population spatial structure, environmental and demographic variation
684 across the population, as well as variation in individual quality within each patch. Starting
685 with the neighbour-modulated fitness of a focal recipient, I then derived the marginal fitness
686 effects for the expression of different kinds of behaviours. In particular, I considered
687 behaviours that are expressed conditionally on an individual's quality, but also behaviours
688 that are expressed conditionally on patch quality as well as unconditionally expressed
689 behaviours.

690

691 I showed how to calculate three key variables that mediate the fitness effects of a behaviour,
692 namely: the frequency of patches in each demographic state; the reproductive value of each
693 individual; and the kin selection coefficients of relatedness. I then obtained the expressions
694 for the reproductive success of each recipient from the demographic transition rates that
695 characterise continuous-time models. Given the fitness effect of a behaviour, the demographic

696 rates, and the expressions of reproductive success, I derived a general formulation of
697 Hamilton's rule for a helping and a dispersal trait and the corresponding inclusive fitness
698 interpretation. Next, I illustrated how this general framework can be used to study and unify a
699 wide range of kin selection problems.

700

701 In a first instance, I derived a general expression for the fitness effect of a social behaviour. In
702 the context of a Wright-Fisher infinite island model, Rousset and Ronce [14] derived a similar
703 expression. However, their marginal fitness expression is narrower in scope and partitioned
704 differently. Like my expression, they have partitioned marginal effects according to whether
705 fitness is achieved through a philopatric or dispersed component, according to the
706 demographic state of patches, and according to whether the marginal fitness effects emerge
707 from changes in the reproductive success of individuals or in the demographic state of the
708 focal patch.

709

710 However, their partition of the marginal fitness also differs from mine in significant ways.
711 First, they have not considered variation in individual quality within patches as captured by
712 my framework. Second, while they focused on changes in the demographic state of patches
713 that emerge from alterations in patch size, my framework also consider other kinds of
714 demographic changes that do not involve modifications in patch size. For instance, within my
715 framework, changes in demography can occur through changes in the age composition of the
716 group or in the infected state of group members. Finally, while their work focuses on the
717 study of the marginal fitness effects, I extracted general formulations of Hamilton's rule for
718 major social behaviours from the marginal fitness effects.

719

720 I established a formal link between the neighbour-modulated approach and Hamilton's
721 inclusive fitness theory in a population with complex demographic and spatial structures. This
722 link is significant for several reasons. First, the empirical testing of hypothesis and the design
723 of experiments often involves an inclusive fitness argument and therefore the adoption of an
724 actor-centric perspective, rather than then the recipient-centric perspective adopted in the
725 neighbour-modulated approach [1,3,4,13,18]. Second, the inclusive fitness perspective
726 provides a common ground that is able to connect not only theoretical and empirical thinking
727 but also theoretical studies that employ different methodologies [18]. In particular, recent
728 years have witnessed the proliferation of different methodologies to determine the intensity of
729 selection operating on social traits, from inclusive fitness to multilevel selection. Hamilton's
730 rule and the inclusive fitness perspective provide a conceptual tool to unify and contrast
731 results obtained from different methods.
732
733 Third, while in abstract terms theory suggests that inclusive fitness is as general as natural
734 selection [13,46], in practice, a large number of studies have contested this idea (reviewed in
735 [46]). My work shows that inclusive fitness holds under a wide range of conditions in the
736 context of populations with complex demographic and spatial structures. Finally, there is the
737 general conviction that the neighbour-modulated approach yields the same results than the
738 inclusive fitness approach [12]. However, practitioners of the neighbour-modulated approach
739 not always provide an inclusive fitness interpretation of the behaviour (e.g. 47,48]). When the
740 inclusive fitness perspective is given, it is frequently inconsistent among different studies,
741 with the interpretations of the behaviour varying significantly across studies (e.g. [49]). In
742 some other cases, an interpretation made in one context breakdowns when other scenarios are
743 considered (e.g. [15,21,22,26]). My study provides an unifying conceptual tool that integrates
744 an increasingly disconnected body of work.

745

746 I derived a general form of Hamilton's rule for a helping trait, under both fecundity and
747 survival effects, and for a dispersal trait. From these general forms of Hamilton's rule, I was
748 able to immediately recover previous results that were obtained using the inclusive fitness
749 method [i.e. 33-35]. In particular, I was able to recover results pertaining to the evolution of a
750 helping trait under fecundity effects [33], the evolution of parasite virulence [34], and the
751 evolution of age-dependent helping and dispersal [35]. The diversity of solutions implicit in
752 the general forms of Hamilton's rule suggests that it should be relatively easy to obtain results
753 for many other social traits, but also that further explorations of my framework is likely to
754 generate deeper insights concerning the fundamental forces driving the evolution of social
755 traits.

756

757 I have established a formal link between stochastic processes and the equations that
758 characterise three key quantities in kin selection models: patch frequencies, relatedness and
759 reproductive value. In all three previous studies, Alizon and Taylor [33], and Wild et al. [34],
760 and Rodrigues [35], the equations that characterise these three key quantities were obtained
761 using heuristic arguments. Above, I was able to provide a formal link between the theory of
762 stochastic processes and the equations for each of these quantities under very general
763 demographic and ecological conditions.

764

765 The general formulation of the problem under consideration can be easily extended to
766 consider cases in which parental quality is correlated with offspring quality. In such cases, the
767 class of the offspring will depend on the quality of the parent, and this has an impact on the
768 selection on helping behaviours. The fidelity of quality inheritance between parents and
769 offspring has an impact in the demographic and genetic structure of the population, which

770 may either promote or inhibit the evolution of cooperation. However, when I consider
771 unconditional phenotypes, the fidelity of quality inheritance has little (if any) influence on the
772 evolution of social behaviour. More generally, the impact of quality inheritance on the
773 evolution of helping has received little attention, and therefore future studies should explore
774 this problem further.

775

776 I focused on an asexually-reproducing and haploid species. However, extending the current
777 framework to sexually reproducing populations is straightforward. This involves additional
778 sex-classes in addition to other types of class-structures. In such cases, the calculation of
779 reproductive value must take sex-structure into account by considering the genetic
780 contribution of individuals to each sex class [16,17]. Moreover, the calculation of the
781 coefficients of consanguinity involves the evaluation of the levels of inbreeding within each
782 sex in addition to the coefficient of consanguinity between social partners [16,17].

783

784 I considered heterogeneity both between and within patches. In the examples provided above,
785 heterogeneity between patches emerges because the size or composition of patches may vary
786 owing to the life history of individuals. In some cases, however, differences in patch quality
787 may be imposed owing to extrinsic forces, such as environmental changes [e.g. 15]. Such
788 cases are easily handled within the framework developed in this study. In addition, I have
789 considered behaviours expressed conditionally on the quality of the actor and on the quality
790 of the recipient. However, we could consider other scenarios, such as when the behaviour is
791 expressed conditionally solely on the quality of the recipients, in which cases one has to
792 define the set of classes that enact the behaviours (e.g. Rodrigues and Gardner [31]).

793

794

795 **Competing interests**

796

797 I have no competing interests.

798

799 **Data Availability**

800

801 This study did not use any data.

802

803 **Competing Interests**

804

805 I have no competing interests.

806

807 **Authors' Contributions**

808

809 AMMR designed and carried out the analysis of the model. AMMR wrote the manuscript.

810 The author gave approval for publication.

811

812 **Funding**

813

814 The author received no funding for this study.

815

816 **Research Ethics**

817

818 This study did not require an ethical assessment.

819

820 **Animal Ethics**

821

822 This study did not require an animal ethical assessment.

823

824 **Permission to Carry out Field Research**

825

826 No permission to carry out field research was required.

827

828 **Acknowledgements**

829

830 I thank Wolfson College, Cambridge for support and Behavioural Ecology Group, Cambridge
831 for helpful discussions.

832

833 **References**

834

835 1. Hamilton WD. 1964. The genetical evolution of social behaviour. I & II. *J Theor Biol* **7**, 1–
836 52.

837

838 2. Hamilton WD. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**,
839 1218–1220.

840

841 3. West SA, Griffin AS, Gardner A. 2007. Evolutionary explanations for cooperation. *Curr*
842 *Biol* **17**, R661–R672.

843

844 4. Bourke AFG. 2011. *Principles of social evolution*. Oxford, UK: Oxford University Press.

845

846 5. Charnov EL. 1977. An elementary treatment of the genetical theory of kin-selection. *J*
847 *Theor Biol* **66**, 541–550.

848

849 6. Frank SA. 1998. *Foundations of social evolution*. Princeton, NJ: Princeton University
850 Press.

851

852 7. West SA, Pen I, Griffin AS. 2002. Cooperation and competition between relatives. *Science*
853 **296**, 72–75.

854

855 8. Lehmann L, Rousset F. 2010. How life history and demography promote or inhibit the
856 evolution of helping behaviours. *Phil Trans R Soc B* **365**, 2599–2617.

857

858 9. Rodrigues AMM, Kokko H. 2016. Models of social evolution: can we do better to predict
859 “who helps whom to achieve what?” *Phil Trans R Soc B* **371**, 20150088.

860

861 10. Taylor PD, Frank SA. 1996. How to make a kin selection model. *J Theor Biol* **180**, 27–37.

862

863 11. Rousset F. 2004. *Genetic structure and selection in subdivided populations*. Princeton,
864 NJ: Princeton University Press.

865

866 12. Taylor PD, Wild G, Gardner A. 2007. Direct fitness or inclusive fitness: how shall we
867 model kin selection? *J Evol Biol* **20**, 301–309.

868

869 13. West SA, Gardner A. 2013. Adaptation and inclusive fitness. *Curr Biol* **23**, R577–84.

- 870
871 14. Rousset F, Ronce O. 2004. Inclusive fitness for traits affecting metapopulation
872 demography. *Theor Popul Biol* **65**, 127–141.
873
- 874 15. Rodrigues AMM, Gardner A. 2012. Evolution of helping and harming in heterogeneous
875 populations. *Evolution* **66**, 2065–2079.
876
- 877 16. Rodrigues AMM, Gardner A. 2015. Simultaneous failure of two sex-allocation invariants:
878 implications for sex-ratio variation within and between populations. *Proc R. Soc B* **282**,
879 20150570.
880
- 881 17. Rodrigues AMM, Gardner A. 2016. The constant philopater hypothesis: a new life history
882 invariant for dispersal evolution. *J Evol Biol* **29**, 153–166.
883
- 884 18. Frank SA. 2013. Natural selection. VII. History and interpretation of kin selection theory.
885 *J Evol Biol* **26**, 1151–1184.
886
- 887 19. Gandon S. 1999. Kin competition, the cost of inbreeding and the evolution of dispersal. *J*
888 *Theor Biol* **200**, 345–364.
889
- 890 20. Pen I. 2000. Reproductive effort in viscous populations. *Evolution* **54**, 293–297.
891
- 892 21. Gardner A, West SA. 2006. Demography, altruism, and the benefits of budding. *J Evol*
893 *Biol* **19**, 1707–1716.
894

- 895 22. Lehmann L, Perrin N, Rousset, F. 2006. Population demography and the evolution of
896 helping behaviors. *Evolution* **60**, 1137–1151.
897
- 898 23. Lehmann L. 2007. The evolution of trans-generational altruism: kin selection meets niche
899 construction. *J Evol Biol* **20**, 181–189.
900
- 901 24. El Mouden C, Gardner A. 2008. Nice natives and mean migrants: the evolution of
902 dispersal-dependent social behaviour in viscous populations. *J Evol Biol* **21**, 1480–1491.
903
- 904 25. Wild G, Fernandes AD. 2009. Investment in the public good through conditional
905 phenotypes of large effect. *J Evol Biol* **22**, 927–941.
906
- 907 26. Rodrigues AMM, Gardner A. 2013. Evolution of helping and harming in viscous
908 populations when group size varies. *Am Nat* **181**, 609–622.
909
- 910 27. Wright S. 1931. Evolution in Mendelian populations. *Genetics* **16**, 97-159.
911
- 912 28. Hamilton WD, May RM. 1977. Dispersal in stable habitats. *Nature* **269**, 578–581.
913
- 914 29. Taylor PD. 1992. Altruism in viscous populations - An inclusive fitness model. *Evol Ecol*
915 **6**, 352–356.
916
- 917 30. Taylor P, Irwin A. 2000. Overlapping generations can promote altruistic behavior.
918 *Evolution* **54**, 1135–1141.
919

- 920 31 Rodrigues AMM, Gardner A. 2013. Evolution of helping and harming in heterogeneous
921 groups. *Evolution* **67**, 2284–2298.
922
- 923 32. Taylor TB, Rodrigues AMM, Gardner A, Buckling A. 2013. The social evolution of
924 dispersal with public goods cooperation. *J Evol Biol* **26**, 2644–2653.
925
- 926 33. Alizon S, Taylor PD. 2008. Empty sites can promote altruistic behavior. *Evolution* **62**,
927 1335–1344.
928
- 929 34. Wild G, Gardner A, West SA. 2009. Adaptation and the evolution of parasite virulence in
930 a connected world. *Nature* **459**, 983–986.
931
- 932 35. Rodrigues AMM. *submitted*. Demography, life history, and the evolution of social
933 behaviour.
934
- 935 36. Taylor PD. 1990. Allele-frequency change in a class-structured population. *Am Nat* **135**,
936 95–106.
937
- 938 37. Otto SP, Day T. 2007. *A biologist's guide to mathematical modeling in ecology and*
939 *evolution*. Princeton, NJ: Princeton University Press.
940
- 941 38. Bulmer MG. 1994. *Theoretical evolutionary ecology*. Sunderland, MA: Sinauer
942 Associates.
943
- 944 39. Moran PAP. 1958. Random processes in genetics. *Proc Camb Phil Soc* **54**, 60–71.

945

946 40. Moran PAP. 1962. *Statistical processes of evolutionary theory*. Oxford, UK: Oxford

947 University Press

948

949 41. Pinsky M, Karlin S. 2011. *An introduction to stochastic modeling*. Burlington, MA:

950 Academic Press, 4th edition.

951

952 42. Allen LJS. 2010. *An introduction to stochastic processes with applications to biology*.

953 Boca Raton, FL: CRC Press.

954

955 43. Fisher RA. 1930. *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.

956

957 44. Grafen A. 2006. A theory of Fisher's reproductive value. *J Math Biol* **53**, 15–60.

958

959 45. Pepper JW. 2000. Relatedness in trait group models of social evolution. *J Theor Biol* **206**,

960 355–368.

961

962 46. Gardner A, West SA, Wild G. 2011. The genetical theory of kin selection. *J Evol Biol* **24**,

963 1020–1043.

964

965 47. Vitalis R, Rousset F, Kobayashi Y, Olivieri I, Gandon S. 2013. The joint evolution of

966 dispersal and dormancy in a metapopulation with local extinctions and kin competition.

967 *Evolution* **67**, 1676–1691.

968

969 48. Leturque H, Rousset F. 2002. Dispersal, kin competition, and the ideal free distribution in
970 a spatially heterogeneous population. *Theor Popul Biol* **62**, 169–180.

971

972 49. Bao M., Wild G. 2012. Reproductive skew can provide a net advantage in both
973 conditional and unconditional social interactions. *Theor Popul Biol* **82**, 200–208.

974

975 **Figure Legends**

976

977 **Figure 1. The directed graph of the embedded Markov chain.** Larger circles represent the
978 focal patch, while smaller circles represent breeding sites. Smaller grey circles represent sites
979 occupied by class-1 individuals, while smaller spotted circles represent sites occupied by
980 class-2 individuals. The indices represent the number of class-1 and class-2 individuals in the
981 patch. Three different types of demographic transitions are represented with the
982 corresponding demographic rates p . First, class-1 and class-2 individuals may give birth to
983 class-1 individuals. Second, class-1 or class-2 individuals may die. Finally, Class-1
984 individuals may become class-2 individuals.

985

986 **Figure 2. Potential for helping (A) as a function of the fidelity of quality inheritance (α).**

987 [A,B] Behaviour conditionally expressed on the quality of actor and recipient. [C]

988 Unconditionally expressed behaviour. Parameter values: $\beta = \alpha, f_1 = 2.2, f_2 = 2.0, \mu_1 = \mu_2 = 0.3,$

989 $d_{10} = d_{20} = 0, d_{11} = d_{12} = d_{22} = 1.$

990

991

992

993

Figure 1.

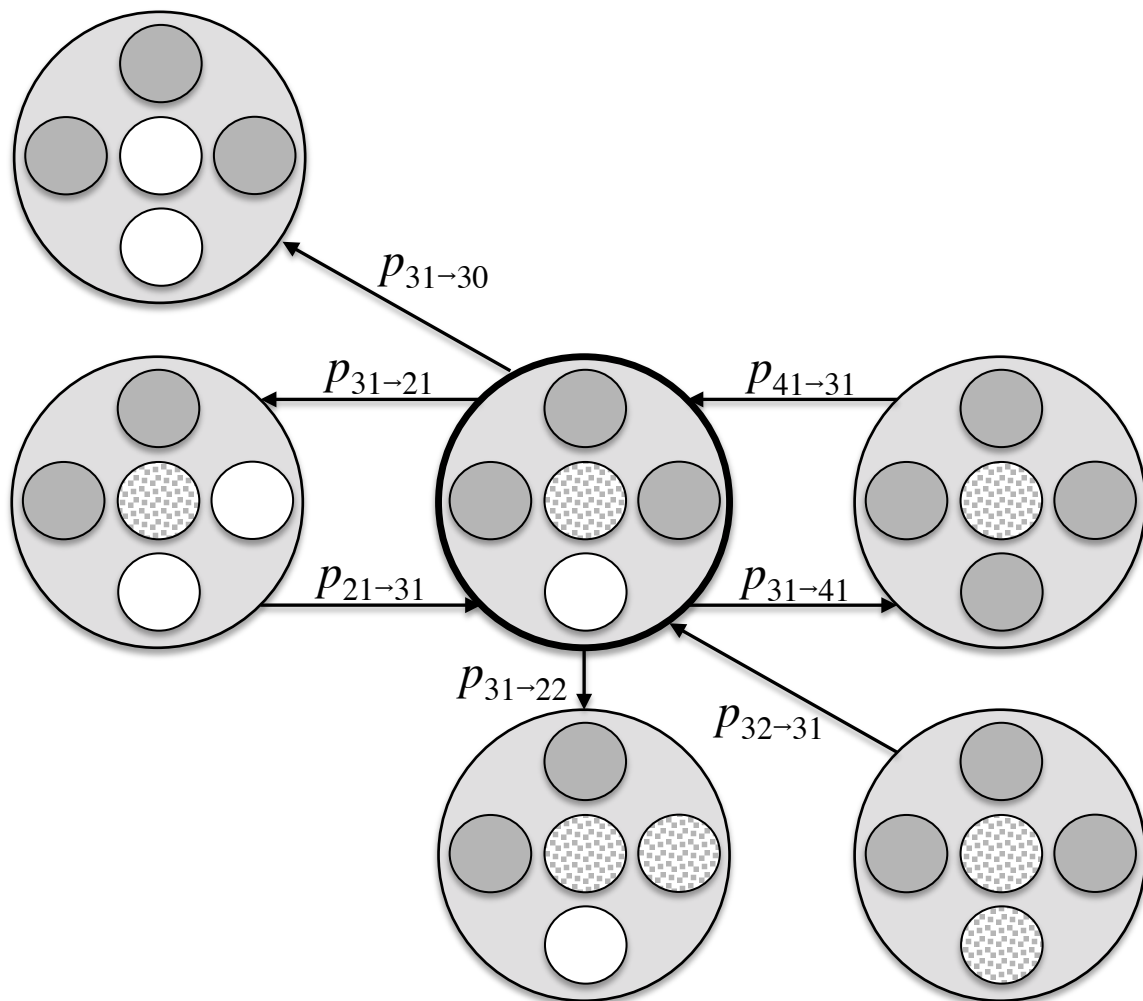


Figure 2.

