Social evolution in structured populations

F. Débarre[1,†], C. Hauert[2] & M. Doebeli[1,2]

Understanding the evolution of social behaviours such as altruism and spite is a long-standing problem that has generated thousands of articles and heated debates. Previous theoretical studies showed that whether altruism and spite evolve may be contingent on seemingly artificial model features, such as which rule is chosen to update the population (for example, birth–death or death–birth), and whether the benefits and costs of sociality affect fecundity or survival. Here we unify these features in a single comprehensive framework. We derive a general condition for social behaviour to be favoured over non-social behaviour, which is applicable in a large class of models for structured populations of fixed size. We recover previous results as special cases, and we are able to evaluate the relative effects of benefits and costs of social interactions on fecundity and survival. Our results highlight the crucial importance of identifying the relative scale at which competition occurs.

[1] Department of Zoology and Biodiversity Research Centre, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia, Canada V6T 1Z4. [2] Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver, British Columbia, Canada V6T 1Z2. † Present address: Centre for Ecology & Conservation, University of Exeter, Cornwall Campus, Tremough, Penryn TR10 9EZ, UK. Correspondence and requests for materials should be addressed to F.D. (email: florence.debarre@normalesup.org).

Why would an individual help or harm others if this reduces its fitness relative to individuals that do not exhibit the social behaviour? Because such behaviours are widespread in nature, numerous theoretical studies have sought to understand the conditions for the evolution of altruism or spite. Although different theoretical frameworks and vocabularies coexist, which sometimes generated heated debates between the proponents of different schools of thought[1–3], they all yield the same kind of conclusions: whether described in terms of games, or direct and indirect fitness or among and within group interactions, altruism requires some form of assortment, so that altruists interact more often with altruists than defectors do[4], while the evolution of spite requires negative assortment[5]. Assortment occurs for instance through clustering in spatially structured (that is, 'viscous') populations[6], or through conditional behaviour when there is kin or type recognition[7–9]. However, the conditions for the evolution of altruism or spite often depend on specific and sometimes artificial model assumptions, such as whether generations are discrete (Wright–Fisher model) or continuous (Moran[10] model)[11,12]. In the latter case, the sequence of events within one time step (death followed by birth, versus birth followed by death) is of crucial importance for the evolutionary outcome[13,14], yet the choice itself seems arbitrary[15]. Also, in most previous models, costs and benefits of altruism or spite (that is, the 'payoffs') are assumed to affect the fecundity of individuals, and costs and benefits for survival have received much less attention[16–18], despite the fact that such effects are equally plausible and should hence be incorporated in general models of social evolution.

Here we present a comprehensive modelling framework that applies to a large class of population structures and not only unites the various assumptions in a single model, but identifies the crucial elements which support the evolution of social behaviour. In structured populations, social behaviour evolves if, for social individuals, the net social benefit of living next to other social individuals outweighs the costs of competing against them. We show that the latter depends on the way the population is updated, the type of social game that is played, and on how social interactions affect individual fecundity and survival.

Results

Dispersal and social structures of the population. We consider a population of fixed size N whose structure is described by two graphs: a dispersal graph $\mathcal{D}$ and an interaction graph $\mathcal{E}$, where each node, also called site, corresponds to one individual of the population. Individuals reproduce clonally, and $\mathcal{D}$ determines dispersal patterns in the population: d_{ij} is the fraction of the offspring produced by the individual at site i that disperse to and compete to colonize site j ($\sum_j d_{ij} = 1$). We assume that the pattern of dispersal is similar for all individuals (technically, $\mathcal{D}$ is transitive, see Supplementary Methods)[14,18,19], and is symmetrical ($d_{ij} = d_{ji}$), and we denote by $d_{\text{self}} = d_{ii}$ the fraction of offspring that remain at their parent's site. Most classical population structures fall into this category[19]: metapopulations, regular lattices, stepping stones, but also groups of groups, and so on (see Fig. 1). Social interactions are reflected in the second graph $\mathcal{E}$, in which e_{ij} measures the strength of the social interaction between the individual at i and the individual at j (scaled so that $\sum_i e_{ij} = 1$). We denote by e_{self} the average strength of social interactions with oneself ($e_{\text{self}} = \sum_i e_{ii}/N$). Finally, the structural average $\overline{\overline{ed}} = \sum_i \sum_i e_{ij} d_{ji}/N$ can be interpreted as the average over all pairs of sites (i, j) of the chance of receiving benefits (e_{ij}) from a site where offspring have been sent (d_{ji}), but also as a measure of the relatedness of an individual to its social
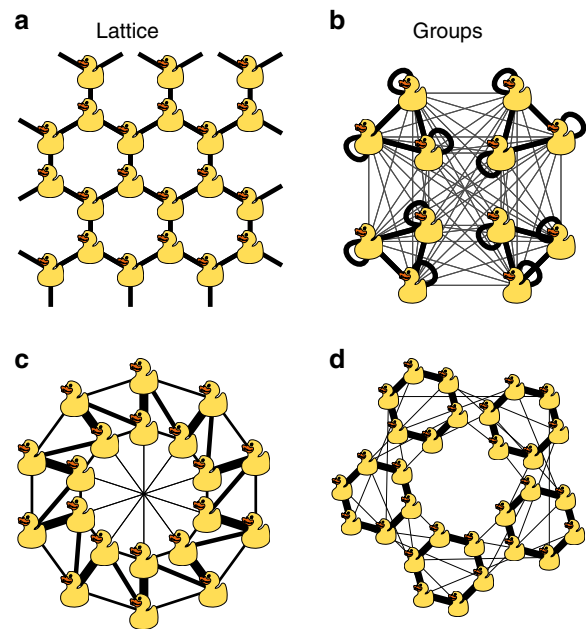

Figure 1 | Transitive dispersal graphs $\mathcal{D}$. Each duck corresponds to one individual living on one site. The thickness of a link between sites i and j is proportional to d_{ij}. (**a**) Lattice: $d_{\text{self}} = 0$; $d_{ij} = 1/k$ between i and its k neighbours, and $d_{ij} = 0$ for non neighbours; here $k = 3$. (**b**) Group-structured population: $d_{ij} = (1-m)/n$ (where m is the chance of dispersing out of the group and n is the size of the groups; here $n = 3$) when i and j belong to the same group (thick links; the loops represent d_{self}); $d_{ij} = m/(N-n)$ when i and j belong to different groups (thin links). Our results are not limited to these two structures and apply to transitive $\mathcal{D}$ graphs in general. (**c**) and (**d**) are examples of other transitive dispersal graphs $\mathcal{D}$.

interaction partners, that is, as a measure of assortment. Hence, four parameters (N, d_{self}, e_{self}, $\overline{\overline{ed}}$) summarize the dispersal and social structure of the population.

A two-step life-cycle and two payoff matrices. The evolution of the population is modelled using a Moran process. Each individual is either social (S) or non-social (NS). Between two time steps, exactly one individual dies and one individual reproduces, so that the size of the population remains N. The identities of the individuals who die and reproduce depend on the individuals' fecundity and survival potential, both being affected by social interactions and by the rules according to which the population is updated. We consider two classic updating rules[13]: death–birth (DB) and birth–death (BD) (these two rules are not restricted to specific population structures, unlike some others, such as budding[20], that are limited to deme-structured populations). In both cases, the first step (death in DB, birth in BD) involves choosing a first individual from all individuals in the population, while the second step (birth in DB, death in BD) involves only those individuals that are connected by dispersal to the site chosen in the first step (dispersal patterns being given by the $\mathcal{D}$ graph). While previous studies only considered effects of social interactions on one of the two steps, the framework we use here allows us to consider the general case where costs and benefits of social interactions can affect both steps, that is, both fecundity and survival. Since effects on fecundity with a BD updating are equivalent to effects on survival with a DB updating in the set of population structures that we consider[19], and *vice versa*, we will give all our heuristic explanations in terms of a DB updating.

The fecundity and survival potential of a given individual living at site i depend on its identity (social or non-social) and on the identities of the individuals it interacts with. To which extent an individual at site j interacts with the individual at i is determined by the e_{ji} term of the $\mathcal{E}$ graph. The effects of this interaction on the recipient's fecundity and survival potential are given by two general payoff matrices:

$$\mathbf{A}_{[1]} = \begin{pmatrix} a_{[1]} & b_{[1]} \\ c_{[1]} & d_{[1]} \end{pmatrix}$$

for the first step (that is, effects on survival with DB), and

$$\mathbf{A}_{[2]} = \begin{pmatrix} a_{[2]} & b_{[2]} \\ c_{[2]} & d_{[2]} \end{pmatrix}$$

for the second step (that is, effects on fecundity with DB).

In both matrices, the a (respectively b) terms refer to payoffs received by a social individual when interacting with another social individual (respectively a non-social), and the d (respectively c) terms refer to payoffs received by non-social individuals, when interacting with a non-social (respectively a social) individual. We use a dynamical system analysis based on moments (singlets, pairs, triplets of individuals and so on) of the distribution of social individuals in a structured population, and we assume weak selection, such that the fitness effects of interactions are small, but individuals are not necessarily phenotypically close[21]. We assume that mutations from one type to the other are rare; a new mutation only occurs after the previous one has been fixed or lost.

The scales of competition. The essential difference between the two steps of the process is the scale of competition (see equation (11), and Supplementary Methods), or, using a kin selection terminology, the identity of the secondary recipients[22] at each step. In a DB updating, an increase in the survival of an individual at site i indirectly harms all the individuals who can send offspring to site i, and the magnitude of this indirect effect on j is determined by d_{ji}. However, an increase in the fecundity of an individual i indirectly harms all other individuals j who would be competing with i for an empty site k, and the magnitude of this indirect effect on j is determined by $\sum_{k=1}^{N} d_{ik} d_{jk}$. Hence, competition in the first step is among all individuals that are one dispersal step away, while competition in the second step is among all individuals that are two dispersal steps away. These two different competition neighbourhoods are illustrated in Supplementary Fig. 1, in the case of a lattice-structured population. In other words, for both DB and BD updating rules, the first step, which involves choosing a first individual globally among all individuals of the population, results in a narrower competitive radius than the second step, in which another individual is chosen locally among the neighbours of the first individual[23]. Thus, whether social interactions affect the first or the second step results in a difference in the spatial scale over which social interactions affect competition. This difference turns out to be crucial for social evolution.

The condition for the evolution of social behaviour. We say that social behaviour evolves when the long-term frequency of social individuals is higher than the frequency of non-social individuals; or, equivalently in the limit of rare and symmetric mutations, when the probability of fixation of initially one social individual in a non-social population is higher than the converse probability of fixation: $\rho_S > \rho_{NS}$. Under our assumptions, we find

that this condition is satisfied when

$$\left\{ \left(1 + e_{\text{self}} - \frac{2}{N}\right)\left(a_{[1]} + a_{[2]} - d_{[1]} - d_{[2]}\right) \right.$$
$$+ \left(1 - e_{\text{self}}\right)\left(b_{[1]} + b_{[2]} - c_{[1]} - c_{[2]}\right)$$
$$+ \left(d_{\text{self}} + \overline{\overline{ed}} - \frac{2}{N}\right)\left(a_{[2]} - d_{[2]}\right) + \left(d_{\text{self}} - \overline{\overline{ed}}\right)\left(b_{[2]} - c_{[2]}\right) \left. \right\} > 0 \quad (1)$$

This condition generalizes previous results[14,18] in two important ways: first, the effects of benefits and costs are not limited to fecundity but may also affect survival at the same time; second, interactions are not restricted to games with equal-gains-from-switching[24] where $a - c = b - d$. Equal-gains-from-switching occurs for instance in the Prisoner's Dilemma (PD, see below) but also, to first order, in any game with small phenotypic differences between individuals[25]. With such payoffs, it is possible to find explicit expressions for ρ_S and ρ_{NS} (see Fig. 2). In contrast, our equation (1) also includes synergistic enhancement or discounting.

Additional insights can be gained by defining an equivalent payoff matrix, $\tilde{\mathbf{A}}$, such that condition (1) can be rewritten as $(1 \ -1) \cdot \tilde{\mathbf{A}} \cdot (1 \ 1)^{\text{T}} > 0$. The equivalent matrix $\tilde{\mathbf{A}}$ can be written as

$$\tilde{\mathbf{A}} = \begin{pmatrix} \sigma_{[1]} a_{[1]} & b_{[1]} \\ c_{[1]} & \sigma_{[1]} d_{[1]} \end{pmatrix} + \xi \begin{pmatrix} \sigma_{[2]} a_{[2]} & b_{[2]} \\ c_{[2]} & \sigma_{[2]} d_{[2]} \end{pmatrix}. \quad (2)$$

For each step i of the life-cycle, $\sigma_{[i]}$ measures the amount of assortment due to the population structure, that is, how much individuals of the same type interact relative to individuals of different types[26] compared with how much they do in large
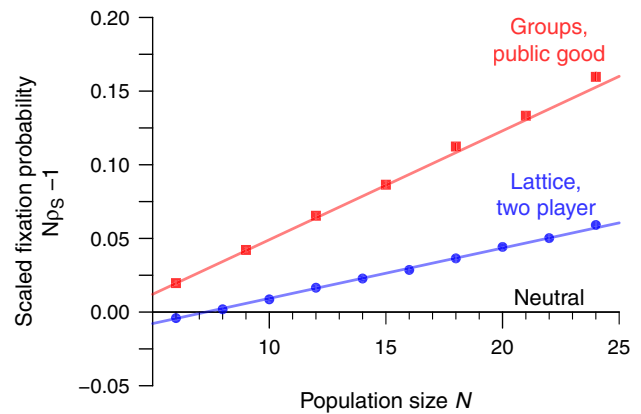

Figure 2 | Fixation probabilities. Comparison of the analytical fixation probability ρ_S with equal-gains-from-switching (solid coloured lines) to the frequency of fixation in numerical simulations (circles and squares), for different values of the total population size N and two classic population structures; results are scaled relative to $1/N$, the neutral expectation $((\rho_S - 1/N)/(1/N) = N\rho_S - 1)$. Two population structures were simulated: in blue, a lattice with $k = 3$ neighbours, two-player games $\left[e_{\text{self}} = d_{\text{self}} = 0, \overline{\overline{ed}} = 1/k\right]$; in red, groups of $n = 3$ individuals with public-good games $\left[e_{\text{self}} = 1/n \text{ and } d_{\text{self}} = (1-m)/n = \overline{\overline{ed}}; m = 0.1\right]$. Fixation probability: denoting $B_{[i]} = c_{[i]} - d_{[i]}$ and $C_{[i]} = d_{[i]} - b_{[i]}$, the approximate fixation probability is $\left[B_{[1]}(N e_{\text{self}} - 1) - C_{[1]}(N-1) + B_{[2]}\left(N e_{\text{self}} + N\overline{\overline{ed}} - 2\right)\right.$ $\left. - C_{[2]}(N + N d_{\text{self}} - 2)\right]/(2N)$. Simulation data: 10^7 runs, $\pm 95\%$ confidence interval (CI) (behind the dots). Parameters: $b_{[1]} = 1\omega$, $c_{[1]} = 0.1\omega$, $b_{[2]} = 10\omega$, $c_{[2]} = 0.5\omega$, $\omega = 0.0025$.

well-mixed populations (in which $\sigma_{[1]} = \sigma_{[2]} = 1$):

$$\sigma_{[1]} = \frac{1 + e_{\text{self}} - 2/N}{1 - e_{\text{self}}},$$

$$\sigma_{[2]} = \frac{1 + e_{\text{self}} + d_{\text{self}} + \overline{\overline{ed}} - 4/N}{1 - e_{\text{self}} + d_{\text{self}} - \overline{\overline{ed}}}. \quad (3)$$

The quantity ξ in equation (2a) measures the relative importance of the second step of the life-cycle, compared with the first:

$$\xi = \frac{1 - e_{\text{self}} + d_{\text{self}} - \overline{\overline{ed}}}{1 - e_{\text{self}}}. \quad (4)$$

In a large well-mixed population, $\xi = 1$.

Contrasting the conditions on each step. By definition of the dispersal graph $\mathcal{D}$, the offspring of an individual are located one dispersal step away, which happens to correspond to the competitive radius during the first step of the Moran process. Individuals are therefore directly competing against their offspring, and the detrimental effects of kin competition exactly cancel the social benefits of living next to related individuals[27,28]. As a result, population structure barely has any effect on the evolution of social behaviour if social interactions affect the first step of the process (see $\sigma_{[1]}$ in equation (3)), compared with large well-mixed

populations. In this case, the relative effect of population structure is limited to the benefits a social individual provides to itself (e_{self}), while small population sizes reduce same-type interactions: $\sigma_{[1]}$ is smaller when N is smaller. This also confirms that even with synergistic effects in the first step ($a_{[1]} - c_{[1]} \neq b_{[1]} - d_{[1]}$), spatial structure does not facilitate the evolution of altruism[29] compared with well-mixed populations. Equation (1) also confirms that in the absence of kin discrimination or synergistic effects ($a_{[1]} - c_{[1]} = b_{[1]} - d_{[1]}$), the evolution of spite ($c_{[1]} - d_{[1]} < 0$) requires small population sizes (small N), and limited self-interactions ($e_{\text{self}} \rightarrow 0$).

In contrast, population structure is of crucial importance for the evolution of social behaviour whenever social interactions affect the second step of the process ($\sigma_{[2]}$ in equation (3) and ξ in equation (4)). This is because the radius of the competitive circle is wider at the second step (two dispersal steps away): individuals are therefore competing against less related individuals, on average, than at the first step[23]. This observation had been made with specific models. For instance, conditions for general games and regular graphs of degree k (and large N) were derived by Ohtsuki *et al.*[13]; these conditions reduced to $(c_{[2]} - d_{[2]})/(d_{[2]} - b_{[2]}) > k$ when $a_{[2]} - c_{[2]} = b_{[2]} - d_{[2]}$. For this restricted set of games, Taylor *et al.*[14] derived a generalization that extended to small populations sizes and covered a broad range of structures, including weighted graphs and distinct interaction $\mathcal{E}$
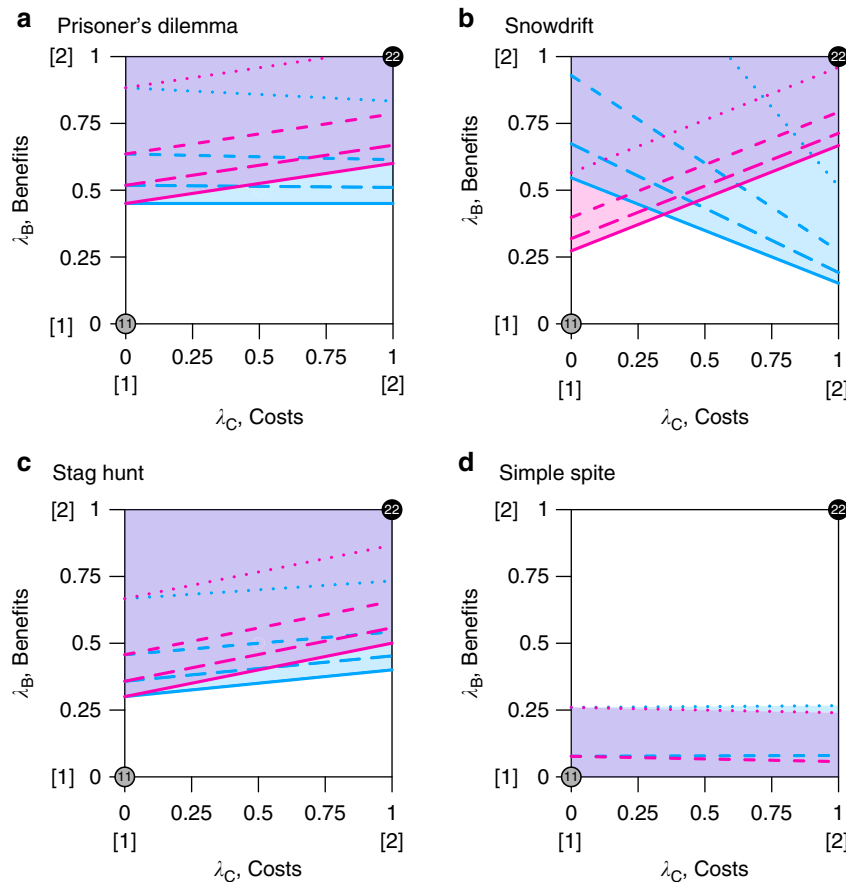

Figure 3 | Evolution of social behaviour under different allocations of benefits and costs. Comparing the conditions for the evolution of social behaviour (altruism in (**a**–**c**), spite in (**d**)), depending on whether benefits and costs affect the first or second step of the process, for four classical games. Results are shown for different population sizes (full line: $N \rightarrow \infty$, large dashed: $N = 60$, dashed: $N = 24$, dotted: $N = 12$) and two structures: in magenta, $d_{\text{self}} = 1/3$, in cyan $d_{\text{self}} = 0$; in both cases $e_{\text{self}} = 0$ and $\overline{\overline{ed}} = 1/3$. Social individuals are favoured ($\rho_S > \rho_{NS}$) in the shaded areas. The grey (benefits and costs on the first step exclusively) and black dots (benefits and costs on the second step exclusively) in the corners correspond to the only parameter combinations that have been analysed previously in this context[13,14,18,19,30]. Parameters: (**a**) Prisoner's dilemma, $c'/b' = 0.15$; (**b**) Snowdrift, $c'/(2b' - c') = 0.65$; (**b**) Stag hunt, $(b' + c')/2b' = 0.6$; (**d**) Simple spite, $c'/b' = 0.02$.

and dispersal $\mathcal{D}$ graphs. In particular, their derivation also included the $(c_{[2]} - d_{[2]})/(d_{[2]} - b_{[2]}) > hg/l$ rule for regular graphs[30], where g and h are the number of neighbours on the g and h graphs, respectively, and l the number of overlapping edges. Our formula (1) not only confirms and generalizes these findings but also offers additional insights, and we illustrate the implications of our results in Fig. 3.

Combining the two steps. We now consider four classical games that can be expressed in terms of benefits and costs (that is, two-parameter games), and we assume that social individuals can allocate these benefits and costs in different proportions to fecundity and survival effects, or more generally to the first and second step in the life-cycle. Then the benefits and costs affecting the first step (fecundity in BD, survival in DB) are $B_{[1]} = b'(1 - \lambda_b)$ and $C_{[1]} = c'(1 - \lambda_c)$, respectively, while the benefits and costs affecting the second step (survival in BD, fecundity in DB) are $B_{[2]} = b'(1 - \lambda_b)$ and $C_{[2]} = c'(1 - \lambda_c)$, respectively. Two of the games that we consider have equal-gains-from-switching: a Prisoner's dilemma (PD) with payoff matrix on step k given by $M_{[k]}^{PD} = \begin{pmatrix} B_{[k]} - C_{[k]} & -C_{[k]} \\ B_{[k]} & 0 \end{pmatrix}$, and the equivalent spite version with negative benefits (SS), whose payoff matrices are $M_{[k]}^{SS} = \begin{pmatrix} -B_{[k]} - C_{[k]} & -C_{[k]} \\ -B_{[k]} & 0 \end{pmatrix}$. We also consider a Snowdrift game (SD) with payoff matrices $M_{[k]}^{SD} = \begin{pmatrix} B_{[k]} - C_{[k]}/2 & -C_{[k]} \\ B_{[k]} & 0 \end{pmatrix}$ and a Stag hunt game (SH) $M_{[k]}^{SH} = \begin{pmatrix} 2B_{[k]} - C_{[k]} & -C_{[k]} \\ B_{[k]} & B_{[k]} \end{pmatrix}$. For each game, we assess how the allocation of benefits and costs on either step of the life-cycle affects the evolution of social behaviour, and do so for two population structures that differ by whether an individual can be replaced by their offspring or not (that is, whether $d_{self} > 0$ or $d_{self} = 0$). The results are illustrated in Fig. 3. We find that the optimal allocation of the benefits only depends on whether these correspond to altruism (positive benefits) or spite (negative benefits). Altruism (Fig. 3a–c) is most favoured if benefits are allocated to the second step of the process, which gives more weight to interactions of individuals of the same type ($\sigma_{[2]} > \sigma_{[1]}$). Spite, on the contrary, is more likely to evolve when the (negative) benefits affect the first step rather than the second one (Fig. 3d), and requires small population sizes. Let us now consider costs. In our examples, the optimal cost allocation depends on both the type of game, the size of the population (N) and on the fraction of an individual's propagules that remain at the exact same site (d_{self}). For example, for a sufficiently large d_{self} (magenta lines in Fig. 3), altruism is most favoured if costs affect the first step of the process. With a low d_{self}, however ($d_{self} = 0$ for the cyan lines in Fig. 3), the allocation of the cost may not matter (for example, Fig. 3a,d, Prisoner's dilemma with large (N), or better be on the second step (for example, Fig. 3b, Snowdrift), or finally better be on the first step (for example, Fig. 3c, Stag hunt).

While we just described cases where social interactions are of the same type on both steps, our framework also allows for the consideration of mixed cases. Focusing on games with equal-gains–from-switching, with $e_{self} = 0$ (like in Fig. 3a,d), our results suggest that, behaviours that are spiteful in the first step and altruistic in the second will be favoured by selection.

Discussion

The simplicity and generality of our condition for the evolution of social behaviour hinges on two standard and widely used

assumptions of weak selection and constant population size. Extending our results to unsaturated populations[17,31], as well as to all possible network structures, is an important future challenge. Meanwhile, our results unify and generalize a great number of existing studies on the evolution of altruism or spite in structured populations, and yield new insights when social behaviour can affect both fecundity and survival. In particular, the results highlight the crucial feature determining the outcome of social evolution: the evolution of social behaviour is determined by the scale at which social interactions affect competition.

Methods

Notation. The population lives in an environment with N sites, labelled $\{1, \ldots, N\}$. An indicator variable $X_i(t)$ gives the occupation of site i at time t: 1 (respectively 0) means that the site is occupied by a social individual (respectively non-social). Notation '$^-$' denotes a population average, and var_S a population variance: $\bar{X}(t) = \sum_{i=1}^{N} X_i(t)/N$ and $\mathrm{var}_S(X(t)) = \sum_{i=1}^{N}(X_i(t) - \bar{X}(t))^2/N = \overline{X^2}(t) - \bar{X}^2(t)$. The index S in the population variance var_S is here to distinguish this variance from the variance of the state of sites ($\mathrm{var}[X_i(t)] = \mathbb{E}_t[X_i^2] - \mathbb{E}_t[X_i]^2$); in the same way there is a distinction between average ($^-$) and expectation ($\mathbb{E}_t[]$). We denote by

$$p_i(t) = \mathbb{E}_t[X_i] = \mathbb{P}(X_i(t) = 1) \tag{5}$$

the expectation of the state of the individual at site i and at time t; a vector $\mathbf{p}(t)$ groups the expected state of all sites, and $p(t) = \overline{p_i(t)}$ is the expected frequency of social individuals in the population at time t. Because of population structure, we also have to take into account the dynamics of pairs and triplets of individuals. We denote by

$$P_{ij}(t) = \mathbb{E}_t[X_i X_j] = \mathbb{P}(X_i(t) = 1, X_j(t) = 1) \tag{6}$$

the expectation of the state of pairs of sites, and group them in a matrix $\mathbf{P}(t)$. We note that the notion of relatedness classically used in kin selection studies is different and involves probabilities of identity in state or probabilities of identity by descent, without conditioning on the type of the individuals: $G_{ij} = \mathbb{P}(X_i(t) = X_j(t))$ Relatedness R_{ij} is often defined as a standardized measure of identity[19]:

$$R_{ij} = \frac{G_{ij} - \overline{\overline{G}}}{1 - \overline{\overline{G}}}, \tag{7}$$

where $\overline{\overline{G}} = \sum_{i=1}^{N} \sum_{j=1}^{N} G_{ij}/N^2$. Our measure of assortment $P_{ij}(t)$ can therefore be expressed as a function of this measure of relatedness R_{ij}:

$$P_{ij}(t) = \mathbb{E}_t[X_i X_j] = (R_{ij} - 1)\mathbb{E}_t[\mathrm{var}_S X] + p(t). \tag{8}$$

We also need to account for associations between triplets of individuals, and we denote by

$$\Pi_{ijk}(t) = \mathbb{E}_t[X_i X_j X_k] = \mathbb{P}(X_i(t) = 1, X_j(t) = 1, X_k(t) = 1) \tag{9}$$

the expectation of the state of triplets of sites, and group them in a three-dimensional array $\mathbf{\Pi}(t)$.

The dispersal and social structures of the population are represented by two graphs $\mathcal{D}$ and $\mathcal{E}$. We denote by d_{ij} the fraction of the offspring of the individual living at site i that is sent to site j, and by e_{ij} the strength of the interaction from i to j; all these parameters are grouped in two matrices $\mathbf{D}$ and $\mathbf{E}$. The dispersal graph is assumed to be symmetric, so that $d_{ij} = d_{ji}$ (or equivalently $\mathbf{D} = \mathbf{D}^T$, where ^{T} denotes transposition); the dispersal graph is also assumed to be transitive, so that the dispersal structure looks the same from every site[14] (note that the transitivity assumption is not required to derive equation (11), but it is needed to obtain explicit expressions for the expected state of pairs of individuals (equation (12))).

The fecundity and survival of individuals in the population are affected by pairwise interactions with other individuals, described by the interaction graph $\mathcal{E}$. For our derivation, it is more convenient to rewrite the payoff matrices $\mathbf{A}_{[1]}$ and $\mathbf{A}_{[2]}$ (whose expressions are given in the main text) as follows:

$$\mathbf{A}_{[i]} = \begin{pmatrix} b'_{[i]} - c'_{[i]} + d'_{[i]} & -c'_{[i]} \\ b'_{[i]} & 0 \end{pmatrix}, \tag{10}$$

with $b'_{[i]} = c_{[i]} - d_{[i]}$, $c'_{[i]} = d_{[i]} - b_{[i]}$, and $d'_{[i]} = a_{[i]} - b_{[i]} - c_{[i]} + d_{[i]}$. The two formulations are equivalent[32]. The effects of pairwise interactions add up, and we assume that their effects on fecundity and survival are weak, of order $\omega \ll 1$: in other words, selection is weak. Note that this type of weak selection differs from weak selection due to small phenotypic differences ('δ-weak selection'[21]) classically used in kin selection models.

Expected change in the frequency of social individuals. We first derive a general equation for the change in the frequency of social individuals, under weak selection and for a symmetric dispersal graph ($\mathbf{D} = \mathbf{D}^T$). We show in the Supplementary Methods how this expression relates to the Price equation[33]. Denoting by $\mathrm{Tr}(\mathbf{M})$ the trace, that is, the sum of diagonal elements, of a matrix $\mathbf{M}$,

the expected change in the frequency of social individuals in the population is (details of the calculations are presented in the Supplementary Methods):

$$\Delta p(t) = \frac{\omega}{N^2}\left[\overbrace{(\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{P})}^{\text{direct effects}}\quad \overbrace{-\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{P}\cdot\mathbf{D}))}^{\text{competition = secondary effects}}b'_{[1]}\right.$$

$$-(\text{Tr}(\mathbf{P})\qquad -\text{Tr}(\mathbf{P}\cdot\mathbf{D}))c'_{[1]}$$
$$+(\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{P})\qquad -\text{Tr}(\mathbf{E}^{\text{T}}\cdot\Pi\cdot\mathbf{D}))d'_{[1]}$$
$$+(\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{P})\qquad -\text{Tr}(\mathbf{E}^{\text{T}}\cdot\Pi\cdot\mathbf{D}\cdot\mathbf{D}))b'_{[2]}$$
$$+(\text{Tr}(\mathbf{P})\qquad -\text{Tr}(\mathbf{P}\cdot\mathbf{D}\cdot\mathbf{D}))c'_{[2]}$$
$$\left.+(\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{P})\qquad -\text{Tr}(\mathbf{E}^{\text{T}}\cdot\Pi\cdot\mathbf{D}\cdot\mathbf{D}))d'_{[2]}\right] \qquad (11)$$

Supplementary Figure 2 (direct effects) and Supplementary Fig. 3 (secondary effects) illustrate the different terms of equation (11). The first three lines correspond to the first step (fecundity effects under BD, survival effects under DB), the last three lines to the second step (survival effects under BD, fecundity effects under DB). For each step, we can distinguish between direct effects (first column) and competition terms (second column). These competition terms correspond to secondary effects in kin selection models[22] or to circles of compensation[23], and differ among the two steps: the competitive radius includes individuals one dispersal step away in the first step ($\mathbf{P}\cdot\mathbf{D}$ terms), and two dispersal steps away in the second step ($\mathbf{P}\cdot\mathbf{D}\cdot\mathbf{D}$ terms). This equation is dynamically not closed, because it depends on higher moments such as pairs ($\mathbf{P}$) and triplets (Π) of social individuals.

Evaluating the moments. We show that the dynamics of pairs and triplets occur at a much faster time scale than the dynamics of the average frequency, so that we can evaluate them using a separation of time scales[34] or quasi-equilibrium approximation.

Using the fact that the dispersal graph $\mathcal{D}$ is transitive[14,23] and that selection is weak, we obtain the following equalities for the pairs ($\mathbf{1}_{N,N}$ is a N-by-N matrix containing only 1s, and $\mathbf{I}_N$ is the identity matrix):

$$\begin{aligned}\mathbf{P}-\mathbf{P}\cdot\mathbf{D} &= -\phi(t)\mathbf{1}_{N,N}+N\phi(t)\mathbf{I}_N,\\ \mathbf{P}-\mathbf{P}\cdot\mathbf{D}\cdot\mathbf{D} &= -2\phi(t)\mathbf{1}_{N,N}+N\phi(t)(\mathbf{I}_N+\mathbf{D}),\end{aligned} \qquad (12)$$

where $\phi(t)=(N/(N-1))\mathbb{E}_t[\text{var}_S\,X]$ (details of the calculations are in the Supplementary Methods). Note that we did not use the pair approximation[35] to derive equation (12).

For triplets of individuals, we use recent results[26,32] showing how to express terms with triplets as functions of pairs plus a frequency-dependent term scaled by a factor $\alpha_{[i]}$ for each step i (see Supplementary Methods for details). The $\alpha_{[i]}$ factors will remain implicit, but they will vanish in the final condition for the evolution of social behaviour.

Explicit dynamics. We use the expressions we derived for pairs (equation (12)) and triplets of individuals back in the equation for the frequency dynamics (equation (11)), thus arriving at a closed dynamical system, given below in equation (16). We denote by $d_{\text{self}}=d_{ii}$, the fraction of propagules that remain at their parent's site (the same for all sites i on a transitive graph), by $e_{\text{self}}=(1/N)\sum_{l=1}^{N}e_{ll}$ the average interaction with oneself, and finally a compound parameter summarizing the dispersal and interaction graphs (recall that $d_{kl}=d_{lk}$),

$$\overline{\overline{ed}}=\frac{1}{N}\sum_{l=1}^{N}\sum_{k=1}^{N}e_{lk}d_{kl}=\frac{1}{N}\sum_{l=1}^{N}\sum_{k=1}^{N}e_{lk}d_{lk}=\frac{1}{N}\text{Tr}(\mathbf{E}^{\text{T}}\cdot\mathbf{D}) \qquad (13)$$

that can be interpreted as the average chance of receiving benefits (e_{lk}) from a site where offspring have been sent (d_{kl}). We then define two compound parameters, $\sigma_{\mathcal{DE}}$ and $\tau_{\mathcal{DE}}$, that depend on the population structure and the different payoffs:

$$\sigma_{\mathcal{DE}}=\omega\left(+b'_{[1]}\frac{Ne_{\text{self}}-1}{N-1}-c'_{[1]}\right.$$
$$+d'_{[1]}\left(\frac{N+Ne_{\text{self}}-2}{2(N-1)}-\frac{\alpha_{[1]}}{2}\right)$$
$$+b'_{[2]}\frac{Ne_{\text{self}}+N\overline{\overline{ed}}-2}{N-1}-c'_{[2]}\frac{N+Nd_{\text{self}}-2}{N-1} \qquad (14)$$
$$\left.+d'_{[2]}\left(\frac{N+Nd_{\text{self}}+Ne_{\text{self}}+N\overline{\overline{ed}}-4}{2(N-1)}-\frac{\alpha_{[2]}}{2}\right)\right),$$

and

$$\tau_{\mathcal{DE}}=\omega\left(d'_{[1]}\alpha_{[1]}+d'_{[2]}\alpha_{[2]}\right). \qquad (15)$$

With these definitions, we find that the expected change in the frequency of social individuals can be written as follows:

$$\Delta p(t)=\underbrace{(\sigma_{\mathcal{DE}}+\tau_{\mathcal{DE}}p(t))}_{s_{\mathcal{DE}}(p(t))}\times\mathbb{E}[\text{var}_S(p(t))] \qquad (16)$$

In other words, the structured population behaves as a well-mixed one, under linear frequency-dependent selection given by $s_{\mathcal{DE}}(p(t))$.

Fixation probabilities. The fixation probability of a single mutant under a Moran process with a linear frequency-dependent selection coefficient is a classical result[36–38]. Accordingly, the fixation probability of a social mutant in a population of non-social individuals, when selection is weak, can be approximated as

$$\rho_S\approx\frac{1}{N}+\frac{N-1}{2N}\left(\sigma_{\mathcal{DE}}+\tau_{\mathcal{DE}}\frac{N+1}{3N}\right) \qquad (17)$$

Reciprocally, the fixation probability of initially one non-social individual in a population of social individuals is

$$\rho_{\text{NS}}\approx\frac{1}{N}+\frac{N-1}{2N}\left(-(\sigma_{\mathcal{DE}}+\tau_{\mathcal{DE}})+\tau_{\mathcal{DE}}\frac{N+1}{3N}\right) \qquad (18)$$

We define the evolutionary success of social individuals by the condition

$$\rho_S>\rho_{\text{NS}} \qquad (19)$$

Using equations (17) and (18), this condition becomes $2\sigma_{\mathcal{DE}}+\tau_{\mathcal{DE}}>0$. Finally, using the definitions (14) and (15) we obtain equation (1) in the main text.

Simulations. Stochastic simulations were coded in C; the population was updated following a Moran process as described in the main text, and the simulation stopped when the social trait was either fixed or lost. Fixation probabilities, with initially only one social individual in the population, were estimated after 10^7 runs for each parameter combination. The simulation scripts are available on Dryad: doi:10.5061/dryad.r28qk.

References

1. Nowak, M. A., Tarnita, C. E. & Wilson, E. O. The evolution of eusociality. *Nature* **466,** 1057–1062 (2010).
2. Abbot, P. *et al.* Inclusive fitness theory and eusociality. *Nature* **471,** E1–E4 (2011).
3. Lion, S., Jansen, V. A. & Day, T. Evolution in structured populations: beyond the kin versus group debate. *Trends. Ecol. Evol.* **26,** 193–201 (2011).
4. Fletcher, J. A. & Doebeli, M. A simple and general explanation for the evolution of altruism. *Proc. R. Soc. B* **276,** 13–19 (2009).
5. Gardner, A. & West, S. A. Spite and the scale of competition. *J. Evol. Biol.* **17,** 1195–1203 (2004).
6. Hamilton, W. The genetical evolution of social behaviour. i. *J. Theor. Biol.* **7,** 1–16 (1964).
7. Jansen, V. A. A. & van Baalen, M. Altruism through beard chromodynamics. *Nature* **440,** 663–666 (2006).
8. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314,** 1560–1563 (2006).
9. Lehmann, L. & Keller, L. The evolution of cooperation and altruism-a general framework and a classification of models. *J. Evol. Biol.* **19,** 1365–1376 (2006).
10. Moran, P. A. P. Random processes in genetics. *Math. Proc. Cambridge Phil. Soc.* **54,** 60–71 (1958).
11. Taylor, P. D. & Irwin, A. J. Overlapping generations can promote altruistic behavior. *Evolution* **54,** 1135–1141 (2000).
12. Lehmann, L., Keller, L. & Sumpter, D. J. T. The evolution of helping and harming on graphs: the return of the inclusive fitness effect. *J. Evol. Biol.* **20,** 2284–2295 (2007).
13. Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441,** 502–505 (2006).
14. Taylor, P. D., Day, T. & Wild, G. Evolution of cooperation in a finite homogeneous graph. *Nature* **447,** 469–472 (2007).
15. Zukewich, J., Kurella, V., Doebeli, M. & Hauert, C. Consolidating birth-death and death-birth processes in structured populations. *PLoS One* **8,** e54639 (2013).
16. Nakamaru, M. & Iwasa, Y. The coevolution of altruism and punishment: role of the selfish punisher. *J. Theor. Biol.* **240,** 475–488 (2006).
17. Lion, S. & Gandon, S. Life history, habitat saturation and the evolution of fecundity and survival altruism. *Evolution* **64,** 1594–1606 (2010).
18. Taylor, P. Birth-death symmetry in the evolution of a social trait. *J. Evol. Biol.* **23,** 2569–2578 (2010).
19. Taylor, P., Lillicrap, T. & Cownden, D. Inclusive fitness analysis on mathematical groups. *Evolution* **65,** 849–859 (2011).
20. Lehmann, L. & Rousset, F. How life history and demography promote or inhibit the evolution of helping behaviours. *Phil. Trans. R. Soc. B* **365,** 2599–2617 (2010).
21. Wild, G. & Traulsen, A. The different limits of weak selection and the evolutionary dynamics of finite populations. *J. Theor. Biol.* **247,** 382–390 (2007).
22. West, S. A. & Gardner, A. Altruism, spite, and greenbeards. *Science* **327,** 1341–1344 (2010).
23. Grafen, A. & Archetti, M. Natural selection of altruism in inelastic viscous homogeneous populations. *J. Theor. Biol.* **252,** 694–710 (2008).

24. Nowak, M. & Sigmund, K. The evolution of stochastic strategies in the prisoner's dilemma. *Acta. Appl. Math.* **20,** 247–265 (1990).

25. Taylor, P. D., Day, T. & Wild, G. From inclusive fitness to fixation probability in homogeneous structured populations. *J. Theor. Biol.* **249,** 101–110 (2007).

26. Tarnita, C. E., Ohtsuki, H., Antal, T., Fu, F. & Nowak, M. A. Strategy selection in structured populations. *J. Theor. Biol.* **259,** 570–581 (2009).

27. Taylor, P. D. Inclusive fitness in a homogeneous environment. *Proc. Rl. Soc. Lond. B* **249,** 299–302 (1992).

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

29. Ohtsuki, H. Does synergy rescue the evolution of cooperation? an analysis for homogeneous populations with non-overlapping generations. *J. Theor. Biol.* **307,** 20–28 (2012).

30. Ohtsuki, H., Nowak, M. A. & Pacheco, J. M. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.* **98,** 108106 (2007).

31. Alizon, S. & Taylor, P. Empty sites can promote altruistic behavior. *Evolution* **62,** 1335–1344 (2008).

32. Taylor, P. & Maciejewski, W. An inclusive fitness analysis of synergistic interactions in structured populations. *Proc. R. Soc. B* **279,** 4596–4603 (2012).

33. Price, G. R. Selection and covariance. *Nature* **227,** 520–521 (1970).

34. Law, R. & Dieckmann, U. *The Geometry of Ecological Interactions: Simplifying Spatial Complexity* 252–270 (Cambridge University Press, 2000).

35. Matsuda, H., Ogita, N., Sasaki, A. & Satō, K. Statistical mechanics of population. *Prog. Theor. Phys.* **88,** 1035–1049 (1992).

36. Nowak, M., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428,** 646–650 (2004).

37. Durrett, R. *Probability Models for DNA Sequence Evolution, vol. XII of Probability and Its Applications* 2nd edn, Ch. 1 (Springer, New York, 2008).

38. Huang, W. & Traulsen, A. Fixation probabilities of random mutants under frequency dependent selection. *J. Theor. Biol.* **263,** 262–268 (2010).

Author contributions

F.D. designed and analysed the model; all authors wrote the manuscript; F.D. wrote the Supplementary Information.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Débarre, F. *et al.* Social evolution in structured populations. *Nat. Commun.* 5:3409 doi: 10.1038/ncomms4409 (2014).