



# Stochastic evolution of staying together

Whan Ghang<sup>a,b</sup>, Martin A. Nowak<sup>a,b,c,\*</sup>

<sup>a</sup> Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA

<sup>b</sup> Department of Mathematics, Harvard University, Cambridge, MA 02138, USA

<sup>c</sup> Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA



## HIGHLIGHTS

- Staying together is a crucial operation for construction of complexity in biology.
- Staying together means that cells do not separate after division.
- We study the evolution of staying together.
- We explore a stochastic process with finite population size.
- We derive exact results for the limit of weak selection.

## ARTICLE INFO

### Article history:

Received 27 March 2014

Received in revised form

9 June 2014

Accepted 20 June 2014

Available online 30 June 2014

### Keywords:

Evolution of complexity

Multi-cellularity

Fixation probability

Finite population

## ABSTRACT

Staying together means that replicating units do not separate after reproduction, but remain attached to each other or in close proximity. Staying together is a driving force for evolution of complexity, including the evolution of multi-cellularity and eusociality. We analyze the fixation probability of a mutant that has the ability to stay together. We assume that the size of the complex affects the reproductive rate of its units and the probability of staying together. We examine the combined effect of natural selection and random drift on the emergence of staying together in a finite sized population. The number of states in the underlying stochastic process is an exponential function of population size. We develop a framework for any intensity of selection and give closed form solutions for special cases. We derive general results for the limit of weak selection.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

This paper is part of the effort to explore how staying together (ST) can contribute to the emergence of complexity in evolution (Tarnita et al., 2013; Olejarz and Nowak, in press). ST means that reproductive units do not separate, but stay together. For example, cells that have divided can remain attached to each other forming multi-cellular filaments or aggregates. ST in the context of cellular division can therefore lead to the evolution of multi-cellularity, which is a major topic of investigation (Bell and Mooers, 1997; Bonner, 1998, 2008; Maynard Smith and Szathmari, 1998; Michod, 1999, 2007; Furusawa and Kaneko, 2000; Carroll, 2001; Pfeiffer and Bonhoeffer, 2003; Kirk, 2003, 2005; King, 2004; Grosberg and Strathmann, 2007; Rainey, 2007; Willensdorfer, 2008; Kolter, 2010; Rossetti et al., 2010, 2011; Koschwanez et al., 2011; Ratcliff et al., 2012, 2013; Norman et al., 2013). Another example of ST is that the offspring of a social insect do not leave the nest but stay with their mother and participate in raising further offspring

(Wilson, 1971; Gadagkar, 1994, 2001; Hunt, 2007; Hölldobler and Wilson, 2009). ST in the context of subsocial insects is a trajectory for the evolution of eusociality (Nowak et al., 2010a). Another example of ST is that reproducing intra-cellular symbionts remain in the same host cell. The evolution of eukarya by endosymbiosis (Margulis, 1981) is a form of staying together (Tarnita et al., 2013). At the dawn of life protocells enable a staying together of RNA sequences that replicate inside them (Chen et al., 2005; Bianconi et al., 2013). It is therefore of great interest to study fundamental aspects of the evolutionary dynamics of staying together. Previous work has focused on deterministic evolutionary dynamics (Tarnita et al., 2013; Olejarz and Nowak, in press). Here we develop a stochastic approach.

We study the fixation of ST in a population of finite size,  $N$ . We introduce a single mutant that has the ability to stay together and calculate the probability that it reaches fixation in a population where the resident type does not stay together.

Our paper is structured as follows. In Section 2, we describe the basic model and key results. In Section 3, we show the underlying mathematical ideas and derivation of our results. In Section 4, we provide a brief summary and outlook for future research. The Appendix contains detailed derivations.

\* Corresponding author.

## 2. Model and key results

We consider a population of constant and finite size,  $N$ . There are two types:  $A$  has the ability to form complexes by staying together (ST), while  $B$  only exists in single units. We use the notation  $A_i$  to describe a complex of size  $i$ . The largest conceivable complex size is given by the size of the population,  $N$ . In this case the entire population would consist of a single complex.

We assume that the rate of reproduction of  $A$  units depends on the size of the complex. An  $A$  unit in a complex of size  $i$  has reproductive rate  $a_i$ . In comparison a  $B$  unit has a fixed reproductive rate, 1, which determines the time scale.

If a unit within  $A_i$  reproduces, there are two possibilities: (i) the new unit can stay with the complex, in which case we obtain a complex that has grown in size,  $A_{i+1}$ ; or (ii) the new unit leaves the complex, in which case we obtain an additional new complex of size one,  $A_1$ . The former happens with probability  $q_i$  while the latter happens with probability  $1 - q_i$ . Thus, both the rate of reproduction and the probability of ST can depend on the size of the complex.

Reproduction in our system is described by the following biological reactions:



In any one time step we choose a random unit for reproduction proportional to fitness and simultaneously we choose a random unit to die. If a unit in a complex  $A_i$  (with  $i \geq 2$ ) dies, then we obtain a complex that is one unit smaller,  $A_{i-1}$ . If an  $A_1$  unit dies then this complex disappears. Similarly if a  $B$  unit dies, the total number of  $B$  units in the whole population decreases by one.

Death in our system is described by the following biological reactions:



In contrast to the previous work on staying together which was based on deterministic equations (Tarnita et al., 2013; Olejarz and Nowak, in press), we do not consider the removal (death) of entire complexes. Instead in our system individual units die to ensure constant population size. This assumption facilitates the analysis of the stochastic process (Fig. 1).

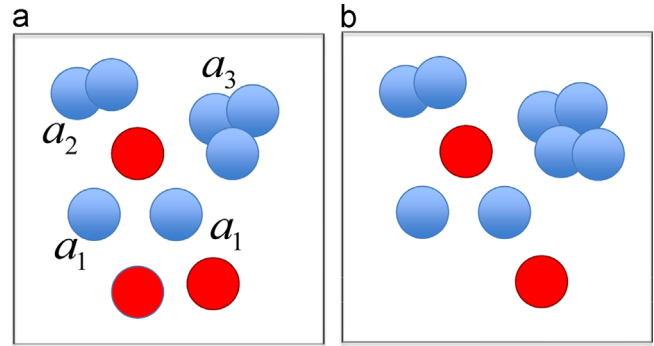
The above notation for the biological reactions in our system is borrowed from chemical kinetics. Note however that we do not use this notation to describe a time continuous process, but a discrete one. Moreover, in our model there is always exactly one birth and one death event to ensure that the population size is strictly constant as in the Moran (1962) process.

Unlike the Moran process (for two types) our system has a very large number of states. If we denote by  $x_i$  the number of complexes of type  $A_i$  and by  $y$  the number of  $B$  units in the population, then a state of the process is given by a vector  $(x_1, x_2, \dots, x_N, y)$  subject to the constraint:

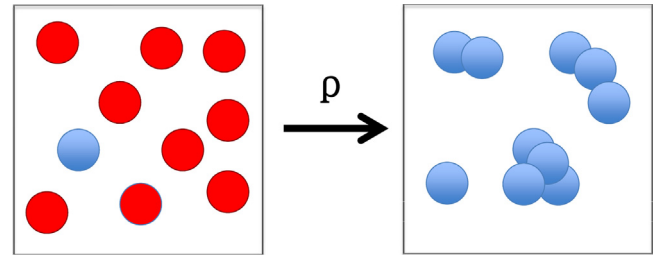
$$y + \sum_{i=1}^N ix_i = N. \tag{3}$$

The total number of states grows exponentially with population size,  $N$ .

The main biological question that we want to answer is the following. We introduce one  $A_1$  unit in a population of  $N - 1$  many  $B$  units and ask what is the probability that the lineage arising



**Fig. 1.** Stochastic evolution of staying together. Blue spheres represent  $A$  units (cells) that have the ability to stay together after reproduction. Red spheres represent  $B$  units which always separate. (a) The fitness values of units depend on their type and for  $A$  units also on the size of the complex in which they are in. An  $A$  unit in a complex of size  $i$  has reproductive rate  $a_i$ .  $B$  units have reproductive rate 1. (b) In any one time step a random unit is chosen for reproduction proportional to fitness and a random unit is chosen for death. The total population size (total number of all units) remains constant. In this example, an  $A$  unit in the complex of size 3 has reproduced; this complex has now size 4. A  $B$  unit has died. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)



**Fig. 2.** We introduce a single  $A$  unit (blue) in a population of  $B$  units (red). The probability that the stochastic processes eventually reach a state where all units in the population are of type  $A$  is called the fixation probability,  $\rho$ , of  $A$ . Note that there is a single absorbing state of only  $B$  units, but there are many states that have the property of consisting of only  $A$  units. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

from this single  $A$  unit takes over the whole population? Thus, we want to calculate the fixation probability of  $A$ . There is either fixation or extinction. Let  $\rho$  be the probability that starting with one  $A$  unit the process eventually reaches a state where all units are of type  $A$  and type  $B$  has become extinct (Fig. 2).

Let us consider a simple case first. If  $q_i$ 's are all 1, then  $A$  units always stay together. Since we start with a single  $A_1$  complex, at any time we have only at most one  $A$  complex. Hence the state space of our stochastic process is simply given by the total number of  $A$  units,  $i = 0, 1, \dots, N$ . There are two absorbing states, 0 and  $N$ , while the remaining states,  $i = 1, \dots, N - 1$ , are transient. Using standard methods (Karlin and Taylor, 1975; Nowak, 2006), we obtain

$$\rho = \left[ 1 + \frac{1}{a_1} + \frac{1}{a_1 a_2} + \dots + \frac{1}{a_1 a_2 \dots a_{N-1}} \right]^{-1}. \tag{4}$$

For our general system we can derive an analytic expression for the fixation probability in the limit of weak selection. In this limit, we assume that the reproductive rate of all units is very close to 1. We introduce the notation  $a_i = 1 + ws_i$ , where  $s_i$  is the payoff (fitness contribution) for being in a complex of size  $i$  and  $w$  is the intensity of selection. For  $w = 0$  we have neutrality. The limit of weak selection is given by  $w \rightarrow 0$ . For studying weak selection, we assume that  $w$  is small enough so that higher order terms of  $w$  are

negligible. In this case, and retaining general  $q_i$ , we obtain

$$\rho = \frac{1}{N} + \frac{N-1}{2N} \frac{\sum_{i=1}^{N-1} C_i s_i}{\sum_{i=1}^{N-1} C_i} w \quad (5)$$

where

$$C_i = \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j). \quad (6)$$

Therefore, the fixation of ST is favored by selection,  $\rho > 1/N$ , if

$$\sum_{i=1}^{N-1} C_i s_i > 0. \quad (7)$$

These results are proved in [Appendices A and B](#).

If the  $q_i$  values are less than one, then the fitness contribution of smaller complexes is more important than that of larger complexes. Since  $C_{i+1}/C_i = (N-1-i)q_i/(N+1-(i+1)q_{i+1})$  which for large  $N$  is approximately  $q_i$ , the coefficient  $(N-1)C_{i+1}/2N\sum_{i=1}^{N-1}C_i$  of  $s_{i+1}$  is  $q_i$  times smaller than the coefficient  $(N-1)C_i/2N\sum_{i=1}^{N-1}C_i$  of  $s_i$ . Hence, the fitness contributions of smaller complexes are more important in determining the fixation of  $A$ .

It is plausible that in many biological settings, ST complexes have a maximum size beyond which further growth is impossible. Let us now consider the case where  $0 < q_i \leq 1$  for all  $i < k$  and  $q_i = 0$  for all  $i \geq k$ . Thus,  $A_k$  is the largest complex that can be generated. There is no growth from  $A_k$  to  $A_{k+1}$  and beyond. For weak selection the probability of fixation is given by

$$\rho = \frac{1}{N} + \frac{N-1}{2N} \frac{\sum_{i=1}^k C_i s_i}{\sum_{i=1}^k C_i} w \quad (8)$$

where

$$C_i = (N+1) \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{k-1} (N+1-jq_j), \quad i = 1, \dots, k-1 \quad (9)$$

and

$$C_k = \prod_{j=1}^{k-1} (N-1-j)q_j \quad (10)$$

For large  $N$ , we have the following asymptotic expression for the fixation probability:

$$\rho = \frac{1}{N} + w \left[ \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right) \right]. \quad (11)$$

Therefore, the fixation of ST with maximum size  $k$  is favored by selection if

$$\sum_{j=1}^k s_j \prod_{i=1}^{j-1} q_i > 0 \quad (12)$$

This condition holds for large population size and weak selection. More generally, we can prove that if (11) holds then for any initial state the fixation of  $A$  is favored by selection and the fixation of  $B$  is opposed by selection.

In the next section, we describe a method how to evaluate the fixation probabilities in our stochastic process.

### 3. Derivation of system of linear equations for fixation probabilities

A state of our stochastic process is given by the pair  $(\vec{x}, y)$ . The vector  $\vec{x} = (x_1, x_2, \dots, x_N)$  denotes the configuration of the  $A$  units in the population. The nonnegative integer  $x_i$  represents the number of complexes of size  $i$ . The largest possible complex has size,  $N$ , which is the total population size. The nonnegative integer  $y$  denotes the number of  $B$  units. Since the total population size is

strictly constant, we have the constraint

$$N = y + \sum_{i=1}^N i x_i \quad (13)$$

We need to calculate the probabilities of all possible transitions from state  $(\vec{x}, y)$ . For this purpose, we introduce notations for birth and death events. The probability that a unit within a complex  $A_i$  is chosen for reproduction is given by

$$b_i = \frac{i a_i x_i}{y + \sum_{j=1}^N j a_j x_j} \quad (14)$$

The probability that a  $B$  unit is chosen for reproduction is given by

$$b = \frac{y}{y + \sum_{j=1}^N j a_j x_j} \quad (15)$$

The probability that a unit within complex  $A_i$  is chosen for death is

$$d_i = \frac{i x_i}{N} \quad (16)$$

The probability that a  $B$  is chosen for death is

$$d = \frac{y}{N} \quad (17)$$

In order to describe transitions in the system, we define the  $i$ th standard unit vector  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$  where the  $i$ th component is 1 and all others are zero. Starting from a generic state,  $(\vec{x}, y)$ , we have the following 10 transitions:

1. A unit in  $A_i$  is chosen for reproduction and its offspring stays with the complex; a  $B$  unit is chosen for death. This event happens with probability  $b_i d_i q_i$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_{i+1} - e_i, y - 1). \quad (18)$$

2. A unit in  $A_i$  is chosen for reproduction and its offspring does not stay; a  $B$  unit is chosen for death. This event happens with probability  $b_i d_i (1 - q_i)$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_1, y - 1). \quad (19)$$

3. A unit in  $A_i$  is chosen for reproduction and its offspring stays with the complex; a unit in the same  $A_i$  complex is chosen for death. This event happens with probability  $b_i (i/N) q_i$ . The state remains the same; the transition is

$$(\vec{x}, y) \rightarrow (\vec{x}, y). \quad (20)$$

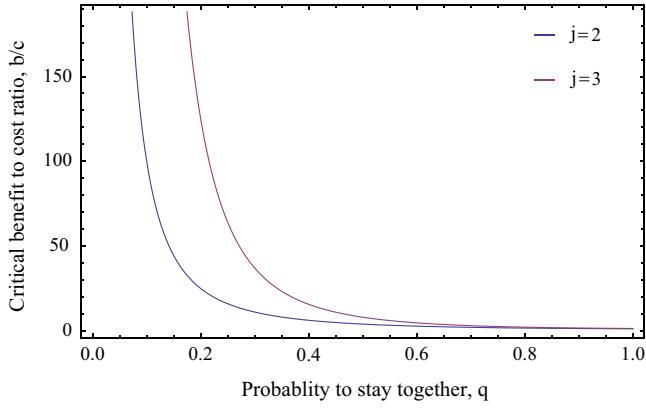
4. A unit in  $A_i$  is chosen for reproduction and its offspring does not stay; a unit in the same  $A_i$  complex is chosen for death. This event happens with probability  $b_i (i/N) (1 - q_i)$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} - e_i + e_{i-1} + e_1, y). \quad (21)$$

5. A unit in  $A_i$  is chosen for reproduction and its offspring stays with the complex; a unit in a different  $A_i$  complex is chosen for death. This event happens with probability  $b_i (d_i - i/N) q_i$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_{i+1} + e_{i-1} - 2e_i, y). \quad (22)$$

6. A unit in  $A_i$  is chosen for reproduction and its offspring does not stay; a unit in different  $A_i$  complex is chosen for death. This event happens with probability  $b_i (d_i - i/N) (1 - q_i)$ .



**Fig. 3.** The critical benefit to cost ratio,  $b/c$ , is shown as a function of the probability to stay together,  $q$ . The fitness landscape is given by  $s_i = -c$  for  $i = 1, 2, \dots, j$  and  $s_i = b - c$  for  $i = j+1, \dots, k$ . The maximum complex size is  $k = 10$ . The minimum size for a beneficial complex is  $j = 2$  (blue) or  $j = 3$  (red). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_1 + e_{i-1} - e_i, y). \quad (23)$$

7. A unit in  $A_i$  is chosen for reproduction and its offspring stays with the complex; a unit in  $A_j$  ( $j \neq i$ ) is chosen for death. This event happens with probability  $b_i d_j q_i$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_{i+1} - e_i - e_j + e_{j-1}, y). \quad (24)$$

8. A unit in  $A_i$  complex is chosen for reproduction and its offspring does not stay; a unit in  $A_j$  ( $j \neq i$ ) is chosen for death. This event happens with the probability  $b_i d_j (1 - q_i)$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} + e_1 - e_j + e_{j-1}, y). \quad (25)$$

9. A  $B$  unit is chosen for reproduction; a unit in  $A_i$  is chosen for death. This event happens with the probability  $b d_i$ . The transition is

$$(\vec{x}, y) \rightarrow (\vec{x} - e_i + e_{i-1}, y + 1). \quad (26)$$

10. A  $B$  unit is chosen for reproduction; a  $B$  unit is chosen for death. This event happens with probability  $b d$ . The state remains the same; the transition is

$$(\vec{x}, y) \rightarrow (\vec{x}, y). \quad (27)$$

As in the standard Moran (1962) process, we choose the birth and death events at the same time to maintain exactly the same number of individuals. Thus, the size of the population is always constant,  $N$ , and the same for both types of events. Note also that if the birth and death events occur in a complex of the same size, then we must distinguish between whether it is exactly the same complex or another complex of the same size, because the two cases yield different transitions (as shown above).

We denote by  $\rho_{\vec{x}, y}$  the probability that  $A$  reaches fixation when starting from state  $(\vec{x}, y)$ .  $A$  has reached fixation, if the population is in a state of the form  $(\vec{z}, 0)$ , which means that  $B$  has become extinct,  $y = 0$ . There are many states where  $A$  has become fixed. All those states fulfill the constraint  $\sum_{i=1}^N i x_i = N$ . The other absorption possibility is that  $A$  becomes extinct and  $B$  reaches

fixation,  $y = N$ . There is exactly one such state, which we denote by  $(\vec{0}, N)$ .

The fixation probabilities are given by the following equation:

$$\begin{aligned} \rho_{\vec{x}, y} = & \sum_i b_i d q_i \rho_{\vec{x} + e_{i+1} - e_i, y-1} \\ & + \sum_i b_i d (1 - q_i) \rho_{\vec{x} + e_i, y-1} \\ & + \sum_i b_i \frac{i}{N} q_i \rho_{\vec{x}, y} \\ & + \sum_i b_i \frac{i}{N} (1 - q_i) \rho_{\vec{x} - e_i + e_{i-1} + e_i, y} \\ & + \sum_i b_i \left( d_i - \frac{i}{N} \right) q_i \rho_{\vec{x} + e_{i+1} + e_{i-1} - 2e_i, y} \\ & + \sum_i b_i \left( d_i - \frac{i}{N} \right) (1 - q_i) \rho_{\vec{x} + e_i + e_{i-1} - e_i, y} \\ & + \sum_i b_i \sum_{j \neq i} d_j q_i \rho_{\vec{x} + e_{i+1} - e_i - e_j + e_{j-1}, y} \\ & + \sum_i b_i \sum_{j \neq i} d_j (1 - q_i) \rho_{\vec{x} + e_i - e_j + e_{j-1}, y} \\ & + \sum_i b d_i \rho_{\vec{x} - e_i + e_{i-1}, y+1} \\ & + b d \rho_{\vec{x}, y} \end{aligned} \quad (28)$$

The summation is always  $i = 1, \dots, N$  unless otherwise stated. The  $\rho$  values satisfy the boundary conditions:  $\rho_{\vec{z}, 0} = 0$  and  $\rho_{\vec{z}, N} = 1$ .

Using the above system of linear equations, we can compute the fixation probabilities, but the problem becomes quite intractable even for moderate  $N$ , because of the large number of states. The number of transient states (with  $0 < y < N$ ) is given by  $p(1) + p(2) + \dots + p(N-1)$ , where  $p(i)$  is the  $i$ th partition number. Hardy and Ramanujan (1918) derived for large  $N$  that

$$p(N) \approx \frac{1}{4\sqrt{3}N} \exp(\pi\sqrt{2N/3}). \quad (29)$$

Using the Stolz–Cesàro Theorem (Stolz, 1885; Cesàro, 1888), we can prove

$$p(1) + p(2) + \dots + p(N) \approx \frac{1}{2\pi\sqrt{2N}} \exp(\pi\sqrt{2N/3}) \quad (30)$$

which is the approximate number of transient states in our process for large  $N$ . For example, if  $N = 100$  there are about  $1.5 \times 10^9$  transient states.

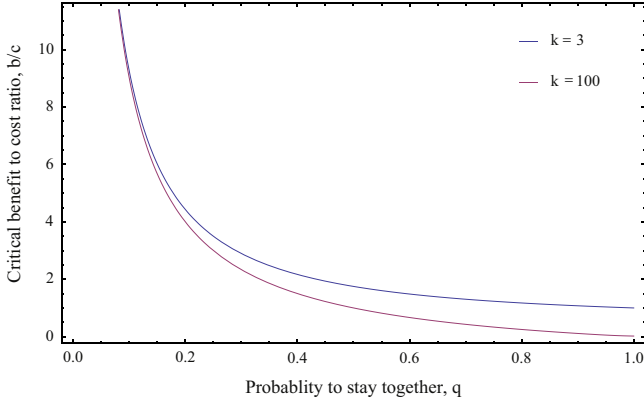
In the Appendix we show how to calculate analytical expressions for the fixation probabilities in the limit of weak selection, thereby providing a derivation of our main results that were listed in Section 2.

#### 4. Conclusion

We have introduced a stochastic process to study the evolution of staying together in populations of finite size. We have derived a special result for the case of any intensity of selection. We have derived general results for the limit of weak selection. Our main result is the following. Suppose that staying together generates complexes up to a maximum size,  $k$ , but not beyond. In this case we have  $q_i > 0$  for  $i = 1, \dots, k-1$  and  $q_i = 0$  for  $i \geq k$ . The reproductive rate of a unit in a complex of size  $i$  is given by  $a_i = 1 + w s_i$ , where  $s_i$  is the selection coefficient and  $w$  is the intensity of selection. For large population size,  $N$ , and in the limit of weak selection,  $w \rightarrow 0$ , we find that staying together is advantageous if

$$s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k > 0. \quad (31)$$

What is the intuition for this result? First of all, we note that we obtain a linear function in the fitness values,  $s_i$ , which is expected given that we consider the limit of weak selection. Second, we observe that the selection coefficients of smaller complexes are



**Fig. 4.** The critical benefit to cost ratio,  $b/c$ , is shown as a function of the probability to stay together,  $q$ . The fitness landscape is given by  $s_i = -c + b(i-1)$  for  $i = 1, \dots, k$ . The maximum complex size is  $k=3$  (blue) or  $k=100$  (red). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

relatively more important than that of larger ones, which makes sense given that larger complexes have to arise from smaller ones and are less abundant in the equilibrium distribution. Specifically, the selection coefficients of large complexes are discounted by the products of the probabilities of staying together describing the sequence of events that would form them. Note also that if all selection coefficients are zero, we have neutrality. If all selection coefficients are positive, then staying together is obviously favored by selection. Interesting conditions emerge, if some selection coefficients are negative while others are positive.

Special cases of the above criterion are

- (i) Let  $q_1 = q_2 = \dots = q_{k-1} = 1$  and  $q_k = 0$ . This means that there is certainty of staying together up to size  $k$ , but not beyond. In this case, ST is advantageous if

$$s_1 + s_2 + \dots + s_k > 0. \tag{32}$$

- (ii) Let  $q_i = q < 1$  and  $q_k = 0$ . This means that there is a fixed probability of staying together, which is the same for all complexes of size smaller than  $k$ . Once size  $k$  is reached all new units leave with certainty. In this case, ST is advantageous if

$$s_1 + qs_2 + \dots + q^{k-1}s_k > 0. \tag{33}$$

We can also consider particular choices for the fitness landscape.

- (iia) Fitness landscape with step function (Fig. 3):  $s_i = -c$  for  $i = 1, 2, \dots, j$  and  $s_i = b - c$  for  $i = j + 1, \dots, k$ . ST entails a cost. Complexes provide a benefit if they are greater than size  $j$ . Here ST is advantageous if

$$\frac{b}{c} > \frac{1 - q^k}{q^j(1 - q^{k-j})}. \tag{34}$$

- (iib) Linear fitness slope (Fig. 4):  $s_i = -c + b(i-1)$ . Again ST entails a cost, but complexes also provide benefits that increase linearly with the size of the complex. Here ST is advantageous if

$$\frac{b}{c} > \frac{(1-q)(1-q^k)}{q(1+(k-1)q^k - kq^{k-1})}. \tag{35}$$

Subsequent work on stochastic staying together might incorporate diffusible public goods (Allen et al., 2013; Olejarz and Nowak, in press), evolutionary game dynamics in populations of

finite size (Nowak et al., 2004; Imhof and Nowak, 2006; Lessard and Ladret, 2007; Ladret and Lessard, 2008), and spatial selection (Nowak and May, 1992; Killingback and Doebeli, 1996, 1998; van Baalen and Rand, 1998; Szabó and Hauert, 2002; Hauert and Doebeli, 2004; Ifti et al., 2004; Santos and Pacheco, 2005; Szabó and Fath, 2007; Gore et al., 2009; Tarnita et al., 2009; Nowak et al., 2010b; Damore and Gore, 2012; Garcia and De Monte, 2012; Broom and Rychtár, 2013; Sanchez and Gore, 2013; Allen and Nowak, 2014). We are also interested in characterizing the mutation-selection distribution of different ST mutants over all possible transient states.

### Appendix A. The limit of weak selection

We can derive analytical results in the limit of weak selection. Let us write the fitness value of an  $A$  unit in a complex  $A_i$  as

$$a_i = 1 + ws_i. \tag{36}$$

The parameter  $w$  scales the intensity of selection. The value  $s_i$  is the selection coefficient that is associated with being in a complex of size  $i$ . As before the fitness value of  $B$  units is 1. The limit of weak selection is given by  $w \rightarrow 0$ . In particular, we can write Taylor expansions keeping linear terms in  $w$  and neglecting higher order terms.

For complete neutrality,  $w=0$ , each unit is equally likely to produce a lineage which will be inherited by the entire population. Therefore, we have

$$\rho_{\vec{x},y} = \frac{N-y}{N}. \tag{37}$$

For the limit of weak selection, adding linear terms in  $w$ , the fixation probabilities must be of the form

$$\rho_{\vec{x},y} \approx \frac{N-y}{N} + wL_{\vec{x},y}. \tag{38}$$

The linear terms,  $L_{\vec{x},y}$ , can be calculated by plugging these fixation probabilities into system (28). We obtain

$$\begin{aligned} N^2 L_{\vec{x},y} = & \frac{y}{N} \sum_i ix_i s_i \\ & + y \sum_i ix_i q_i L_{\vec{x} + e_{i+1} - e_{i,y} - 1} \\ & + y \sum_i ix_i (1 - q_i) L_{\vec{x} + e_{i,y} - 1} \\ & + \sum_i i^2 x_i q_i L_{\vec{x},y} \\ & + \sum_i i^2 x_i (1 - q_i) L_{\vec{x} - e_i + e_{i-1} + e_{i,y}} \\ & + \sum_i i^2 x_i (x_i - 1) q_i L_{\vec{x} + e_{i+1} + e_{i-1} - 2e_{i,y}} \\ & + \sum_i i^2 x_i (x_i - 1) (1 - q_i) L_{\vec{x} + e_{i+1} + e_{i-1} - e_{i,y}} \\ & + \sum_i ix_i q_i \sum_{j \neq i} jx_j L_{\vec{x} + e_{i+1} - e_i - e_j + e_{j-1,y}} \\ & + \sum_i ix_i (1 - q_i) \sum_{j \neq i} jx_j L_{\vec{x} + e_{i+1} - e_j + e_{j-1,y}} \\ & + y \sum_i ix_i L_{\vec{x} - e_i + e_{i-1,y} + 1} \\ & + y^2 L_{\vec{x},y} \end{aligned} \tag{39}$$

The  $L$  terms satisfy the boundary conditions  $L_{\vec{0},N} = 0$  and  $L_{\vec{x},0} = 1$  for all vector  $\vec{x}$ 's. Although the system still contains exponentially many variables, we can solve it explicitly. Let us introduce the sequence of numbers  $u_1, u_2, \dots, u_{N-1}$ . We start with

the definition

$$u_i = \frac{\sum_{i=1}^{N-1} \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j)s_i}{2N \sum_{i=1}^{N-1} \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j)} \quad (40)$$

We define  $u_{i+1}$  for  $i = 1, \dots, N-2$  inductively as follows:

$$q_i(N-i-1)u_{i+1} = (N-iq_i+1)u_i - (N-1)(1-q_i)u_1 - s_i/N \quad (41)$$

We introduce the summation sequence  $v_i = u_1 + \dots + u_i$  for  $i = 1, \dots, N-1$ . Then we construct the vector  $\vec{v}$  whose  $i$ th coordinate is  $v_i$ . Now we are ready to present the exact solution as

$$L_{\vec{x}, y} = y \vec{x} \cdot \vec{v}. \quad (42)$$

This result can be proven by direct computation. See Appendix C for details.

The fixation probability of a single  $A$  unit is

$$\rho = \rho_{e_1, N-1} = \frac{1}{N} + wL_{e_1, N-1}. \quad (43)$$

Remember that  $e_1$  is the  $N$ -dimensional vector  $(1, 0, \dots, 0)$ . We obtain

$$\rho = \frac{1}{N} + \frac{N-1}{2N} \frac{\sum_{i=1}^{N-1} C_i s_i}{\sum_{i=1}^{N-1} C_i} \quad (44)$$

where

$$C_i = \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j). \quad (45)$$

This result holds for the limit of weak selection. Note that the contribution of the selection coefficient  $s_i$  to the fixation probability,  $\rho$ , is  $[N+1-(i+1)q_{i+1}]/[(N-1-i)q_i]$  times larger than that of  $s_{i+1}$ . For large  $N$  the impact of  $s_i$  is  $1/q_i$  times larger than the impact of  $s_{i+1}$ . Thus, for  $q_i$  values less than one, we find that the fitness contributions of smaller complexes are more important than that of larger complexes, which makes sense.

### Appendix B. Weak selection and maximum complex size

We now introduce a maximum complex size,  $k$ . It is possible to grow up to  $A_k$ , but not beyond. If a unit in  $A_k$  is chosen for reproduction, the offspring will always leave. Therefore  $q_i > 0$  for all  $i = 1, \dots, k-1$ , but  $q_k = 0$ . A state of our system is again given by  $(\vec{x}, y)$ , but the vector  $\vec{x}$  has only  $k$  many components  $(x_1, x_2, \dots, x_k)$ . We derive a simple criterion for  $A$  to be advantageous or disadvantageous, which holds for large population size,  $N$ . Consider the quantity

$$C = \sum_{j=1}^k s_j \prod_{i=1}^{j-1} q_i. \quad (46)$$

If  $C > 0$  then for any initial state  $(\vec{x}, y)$  we have  $\rho_{(\vec{x}, y)} > (N-y)/N$ , which means that the fixation of  $A$  is favored by selection starting from any initial state. If  $C < 0$  then for any initial state  $(\vec{x}, y)$  we have  $\rho_{(\vec{x}, y)} < (N-y)/N$ , which means that the fixation of  $A$  is opposed by selection starting from any initial state.

Using Eq. (44) for  $\rho_{e_1, N-1}$ , we have

$$\rho_{e_1, N-1} \approx \frac{1}{N} + \frac{N-1}{2N} \frac{C_1 s_1 + C_2 s_2 + \dots + C_k s_k}{C_1 + C_2 + \dots + C_k} w \quad (47)$$

where

$$C_i = (N+1) \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{k-1} (N+1-jq_j), \quad \forall i = 1, \dots, k-1 \quad (48)$$

and

$$C_k = (N-2)(N-3)\dots(N-k)q_1 q_2 \dots q_{k-1}. \quad (49)$$

Furthermore, we can derive

$$\rho_{\vec{x}, y} = \frac{N-y}{N} + \frac{y(N-y)}{N} \left[ \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1+q_1+q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right) \right] w \quad (50)$$

for any transient initial state  $(\vec{x}, y)$ . The proof is presented in Appendix D.

### Appendix C. Proof of formula $L_{\vec{x}, y} = y \vec{x} \cdot \vec{v}$

**Lemma C.1.** The system of linear equations

$$\frac{1}{N} s_i + q_i(N-i-1)u_{i+1} - (N-iq_i+1)u_i + (N-1)(1-q_i)u_1 = 0 \quad (51)$$

with  $i = 1, \dots, N-1$  has a unique solution. For convention, we assume  $u_N = 0$ .

**Proof.** We show that the corresponding matrix is non-singular. Let  $\vec{t}$  be an arbitrary null vector of the matrix. In other words, the components  $t_i$  satisfy

$$q_i(N-i-1)t_{i+1} - (N-iq_i+1)t_i + (N-1)(1-q_i)t_1 = 0 \quad (52)$$

for  $i = 1, \dots, N-1$ . Let  $M$  be the maximum modulus of  $\vec{t}$  and it is attained at  $1 \leq j \leq N-1$ . Hence  $M = |t_j|$ . By the triangle inequality, we have

$$\begin{aligned} (N-jq_j+1)M &= (N-iq_i+1)|u_j| \\ &= |q_j(N-j-1)t_{j+1} + (N-1)(1-q_j)t_1| \\ &\leq |q_j(N-j-1)t_{j+1}| + |(N-1)(1-q_j)t_1| \\ &\leq q_j(N-j-1)M + (N-1)(1-q_j)M \\ &= (N-jq_j-1)M. \end{aligned}$$

Therefore  $M=0$  and  $t_i$ 's are all zeros. Therefore  $\vec{0}$  is the only null vector of the corresponding matrix and hence the matrix is non-singular, as desired.  $\square$

**Lemma C.2.** In the solution of the above system in C.1, we have

$$u_1 = \frac{\sum_{i=1}^{N-1} \prod_{j=i+1}^{N-1} (N+1-jq_j) \prod_{j=1}^{i-1} (N-1-j)q_j s_i}{2N \sum_{i=1}^{N-1} \prod_{j=i+1}^{N-1} (N+1-jq_j) \prod_{j=1}^{i-1} (N-1-j)q_j} \quad (53)$$

**Proof.** This can be proved by observing the following identity:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N-1} \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j)s_j \\ = \sum_{i=1}^{N-1} \prod_{j=1}^{i-1} (N-1-j)q_j \prod_{j=i+1}^{N-1} (N+1-jq_j)((N-iq_i+1)u_i \\ - q_i(N-i-1)u_{i+1} - (N-1)(1-q_i)u_1) \end{aligned}$$

Note that all the  $u_i$  terms are cancelled for  $i \geq 2$ .  $\square$

**Theorem C.3.** For each population  $N$ , there exists a  $\mathbb{R}^{N-1}$  vector  $\vec{v}$  which only depends linearly on  $s_i$ 's and depends on  $N$  and  $q_i$ 's such that

$$L_{\vec{x}, y} = y \vec{x} \cdot \vec{v} \quad (54)$$

for any population  $N$ , i.e.  $y + \sum_{i=1}^{N-1} ix_i = N$ . The  $v_i$ 's are given by  $v_i = u_1 + u_2 + \dots + u_i$  for each  $i = 1, \dots, N-1$  where  $u_i$ 's are solution of the system of linear equations in C.1.

**Proof.** We need to show that  $L_{\vec{x}, y} = y \vec{x} \cdot \vec{v}$  is a solution by directly plugging it into (39). After cancelling  $\vec{x} \cdot \vec{v}$  terms, we

need to check the following equation:

$$\begin{aligned} & \frac{y}{N} \sum_i i x_i s_i + y(y-1) \sum_i i x_i q_i u_{i+1} \\ & + y(y-1) \sum_i i x_i (1-q_i) u_i \\ & + y \sum_i i^2 x_i (1-q_i) (-u_i + u_1) \\ & + y \sum_i i^2 x_i (x_i - 1) q_i (u_{i+1} - u_i) \\ & + y \sum_i i^2 x_i (x_i - 1) (1-q_i) (u_1 - u_i) \\ & + y \sum_i i x_i q_i \sum_{j \neq i} j x_j (u_{i+1} - u_j) \\ & + y \sum_i i x_i (1-q_i) \sum_{j \neq i} j x_j (u_1 - u_j) \\ & + (y+1) y \sum_i i x_i (-u_i) \\ & = 0 \end{aligned}$$

After dividing both sides by  $y$ , we can organize the left-hand side into

$$\begin{aligned} & \frac{1}{N} \sum_i i x_i s_i + \sum_i i x_i u_{i+1} q_i \left( (y-1) + \sum_{j \neq i} j x_j + i(x_i - 1) \right) \\ & + \sum_i i x_i u_1 (1-q_i) \left( (y-1) + i + i(x_i - 1) + \sum_{j \neq i} j x_j \right) \\ & + \sum_i i x_i u_i \left( -i(1-q_i) - i(x_i - 1) q_i - i(x_i - 1)(1-q_i) - \sum_{j \neq i} j x_j - (y+1) \right) \\ & = \frac{1}{N} \sum_i i x_i s_i + \sum_i i x_i u_{i+1} (N-i-1) q_i \\ & + \sum_i i x_i u_1 (1-q_i) (N-1) - \sum_i i x_i u_i (N-i q_i + 1) \\ & = \sum_i i x_i \left( \frac{1}{N} s_i + q_i (N-i-1) u_{i+1} - (N+1-i q_i) u_i + (N-1)(1-q_i) u_1 \right). \end{aligned}$$

By the definition of  $u_i$ , this is zero as desired. □

**Appendix D. Staying together with maximum complex size,  $A_k$**

The maximum size of staying together is given by  $k$ . We have  $q_i > 0$  for  $i = 1, \dots, k-1$  and  $q_k = 0$ . A state of the process is given by  $(\vec{x}, y)$  where the vector  $\vec{x}$  has  $k$  components,  $(x_1, \dots, x_k)$ .

**Lemma D.1.** For  $i = 1, \dots, k$ , we have

$$Nu_i = \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right). \tag{55}$$

**Proof.** We prove the assertion by induction on  $i$ . If  $i = 1$ , by C.2, we have

$$u_1 = \frac{\sum_{i=1}^k \prod_{j=i+1}^k (N+1-jq_j) \prod_{j=1}^{i-1} (N-1-j) q_j s_j}{2N \sum_{i=1}^k \prod_{j=i+1}^k (N+1-jq_j) \prod_{j=1}^{i-1} (N-1-j) q_j}. \tag{56}$$

Therefore  $Nu_1$  is a rational function of  $N$  such that both numerator and denominator are polynomials of degree  $k-1$ . Hence by comparing the leading coefficients of numerator and denominator, we have the assertion trivially in this case.

Let us now prove the assertion for  $i+1$  ( $1 \leq i \leq k-1$ ) with the induction hypothesis that the assertion is true for  $1, \dots, i$ . We denote

$$\alpha = \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})}. \tag{57}$$

Using C.1, we have

$$Nu_{i+1} - \alpha = \frac{(N-iq_i+1)}{q_i(N-i-1)} (Nu_i - \alpha) - \frac{(N-1)(1-q_i)}{q_i(N-i-1)} (Nu_1 - \alpha) + \frac{2\alpha - s_1}{q_i(N-i-1)}. \tag{58}$$

By the induction hypothesis the right-hand side of the above equation is  $O(1/N)$ . Hence the conclusion follows. □

**Theorem D.2.** Suppose that the maximum complex size is  $k < N$ . We have  $q_i > 0$  for  $i = 1, \dots, k-1$  and  $q_k = 0$ . Then for sufficiently large population size  $N$ , the fixation probability  $A$  is approximately

$$\frac{1}{N} + \left[ \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right) \right] w \tag{59}$$

**Proof.** The assertion follows from direct application of C.3 and D.1. □

**Theorem D.3.** For any state  $(\vec{x}, y)$  with  $0 < y < N$ , we have

$$L_{\vec{x}, y} = \frac{y(N-y)}{N} \left[ \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right) \right]. \tag{60}$$

**Proof.** We denote

$$\alpha = \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})}. \tag{61}$$

By D.1 and definitions of  $v_i$  in C.3, we have that

$$Nv_i = i\alpha + O\left(\frac{1}{N}\right) \tag{62}$$

for each  $i = 1, \dots, k$ . Then we have the following inequality:

$$\begin{aligned} |L_{\vec{x}, y} - \frac{y(N-y)}{N} \alpha| &= y \left| \sum_{i=1}^k x_i \left( v_i - \frac{i\alpha}{N} \right) \right| \\ &= y \left| \sum_{i=1}^k \frac{x_i}{N} (Nv_i - i\alpha) \right| \\ &\leq y \sum_{i=1}^k \frac{x_i}{N} \max_{1 \leq j \leq k} (Nv_j - j\alpha) \\ &= \frac{y(N-y)}{N} \max_{1 \leq j \leq k} |Nv_j - j\alpha| \\ &= \frac{y(N-y)}{N} O\left(\frac{1}{N}\right). \end{aligned}$$

Therefore,  $L_{\vec{x}, y} = (y(N-y)/N)(\alpha + O(1/N))$  as desired. □

As a direct consequence of the above formula, we can calculate the stability of an all  $A$  population against invasion of a single  $B$ . Since the stability is the average of the  $\rho_{\vec{x}, 1}$  for all  $(N-1)$  population distribution of  $A$ 's with the weights according to stationary distribution, the extinction probability of a single  $B$  is

$$\frac{N-1}{N} + \left[ \frac{s_1 + q_1 s_2 + q_1 q_2 s_3 + \dots + q_1 q_2 \dots q_{k-1} s_k}{2(1 + q_1 + q_1 q_2 + \dots + q_1 q_2 \dots q_{k-1})} + O\left(\frac{1}{N}\right) \right] w. \tag{63}$$

**References**

Allen, B., Gore, J., Nowak, M.A., 2013. Spatial dilemmas of diffusible public goods. eLife e01169.  
 Allen, B., Nowak, M.A., 2014. Games on graphs. EMS Surv. Math. Sci. 1 (1), 113–151.  
 Bell, G., Mooers, A.O., 1997. Size and complexity among multicellular organisms. Biol. J. Linn. Soc. 60, 345–363.  
 Bianconi, G., Zhao, K., Chen, I.A., Nowak, M.A., 2013. Selection for replicases in protocells. PLoS Comput. Biol. 9 (5), e1003051.  
 Bonner, J.T., 1998. The origins of multicellularity. Integr. Biol. 1, 27–36.  
 Bonner, J.T., 2008. The Social Amoebae: The Biology of Cellular Slime Molds. Princeton University Press, Princeton.  
 Broom, M., Rychtár, J., 2013. Game-Theoretical Models in Biology. Chapman and Hall/CRC Press, Boca Raton, FL.

- Carroll, S.B., 2001. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* 409, 1102–1109.
- Cesàro, E., 1888. Sur la convergence des séries. *Nouvelles annales de mathématiques Series 3*, 7, 49–59.
- Chen, I.A., Salehi-Ashtiani, K., Szostak, J.W., 2005. RNA catalysis in model protocell vesicles. *J. Am. Chem. Soc.* 127, 13213–13219.
- Damore, J.A., Gore, J., 2012. Understanding microbial cooperation. *J. Theor. Biol.* 299, 31–41.
- Furusawa, C., Kaneko, K., 2000. Complex organization in multicellularity as a necessity in evolution. *Artif. Life* 6, 265–281.
- Gadagkar, R., 1994. Social insects and social amoebae. *J. Biosci.* 19, 219–245.
- Gadagkar, R., 2001. *The Social Biology of Ropalidia marginata: Toward Understanding the Evolution of Eusociality*. Harvard University Press, Cambridge.
- Garcia, T., De Monte, S., 2012. Group formation and the evolution of sociality. *Evolution* 67, 131–141.
- Gore, J., Youk, H., Van Oudenaarden, A., 2009. Snowdrift game dynamics and facultative cheating in yeast. *Nature* 459, 253–256.
- Grosberg, R.K., Strathmann, R.R., 2007. The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Syst.* 38, 621–654.
- Hardy, G.H., Ramanujan, S., 1918. Asymptotic formulae in combinatory analysis. *Proc. Lond. Math. Soc.* 17, 75–115.
- Hauert, C., Doebeli, M., 2004. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428, 643–646.
- Hölldobler, B., Wilson, E.O., 2009. *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. W.W. Norton, London.
- Hunt, J.H., 2007. *The Evolution of Social Wasps*. Oxford University Press, New York.
- Ifti, M., Killingback, T., Doebeli, M., 2004. Effects of neighbourhood size and connectivity on the spatial Continuous Prisoner's Dilemma. *J. Theor. Biol.* 231, 97–106.
- Imhof, L.A., Nowak, M.A., 2006. Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* 52 (5), 667–681.
- Karlin, S., Taylor, H., 1975. *A First Course in Stochastic Processes*. Academic Press, San Diego.
- Killingback, T., Doebeli, M., 1996. Spatial evolutionary game theory: Hawks and Doves revisited. *Proc. R. Soc. B.* 263, 1135–1144.
- Killingback, T., Doebeli, M., 1998. Self-organized criticality in spatial evolutionary game theory. *J. Theor. Biol.* 191, 335–340.
- King, N., 2004. The unicellular ancestry of animal development. *Dev. Cell* 7, 313–325.
- Kirk, D.L., 2003. Seeking the ultimate and proximate causes of Volvox multicellularity and cellular differentiation. *Integr. Comput. Biol.* 43, 247–253.
- Kirk, D.L., 2005. A twelve-step program for evolving multicellularity and a division of labor. *BioEssays* 27, 299–310.
- Kolter, R., 2010. Biofilms in lab and nature: a molecular geneticist's voyage to microbial ecology. *Int. Microbiol.* 13, 1–7.
- Koschwanez, J.H., Foster, K.R., Murray, A.W., 2011. Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. *PLoS Biol.* 9 (8), e1001122.
- Ladret, V., Lessard, S., 2008. Evolutionary game dynamics in a finite asymmetric two-deme population and emergence of cooperation. *J. Theor. Biol.* 255 (1), 137–151.
- Lessard, S., Ladret, V., 2007. The probability of fixation of a single mutant in an exchangeable selection model. *J. Math. Biol.* 54 (5), 721–744.
- Margulis, L., 1981. *Symbiosis in Cell Evolution. Life and its Environment on the Early Earth*. Freeman, San Francisco.
- Maynard Smith, J., Szathmari, E., 1998. *The Major Transition in Evolution*. Oxford University Press, Oxford.
- Michod, R.E., 1999. *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton University Press, Princeton.
- Michod, R.E., 2007. Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl. Acad. Sci. USA* 104, 8613–8618.
- Moran, P.A.P., 1962. *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford.
- Norman, T.M., Lord, N.D., Paulsson, J., Losick, R., 2013. Memory and modularity in cell-fate decision making. *Nature* 503, 481–486.
- Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. *Nature* 359, 826–829.
- Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428, 646–650.
- Nowak, M.A., 2006. *Evolutionary Dynamics*. Harvard University Press, Cambridge, Massachusetts, USA.
- Nowak, M.A., Tarnita, C.E., Wilson, E.O., 2010a. The evolution of eusociality. *Nature* 466, 1057–1062.
- Nowak, M.A., Tarnita, C.E., Antal, T., 2010b. Evolutionary dynamics in structured populations. *Phil. Trans. R. Soc. B* 365, 19–30.
- Olejarz, J., Nowak, M.A., Evolution of staying together in the context of diffusible public goods. *J. Theor. Biol.*, <http://dx.doi.org/10.1016/j.jtbi.2014.06.023>, in press.
- Pfeiffer, T., Bonhoeffer, S., 2003. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proc. Natl. Acad. Sci.* 100, 1095–1098.
- Rainey, P.B., 2007. Unity from conflict. *Nature* 446, 616.
- Ratcliff, W.C., Denison, R.F., Borrello, M., Travisano, M., 2012. Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci.* 109 (5), 1595–1600.
- Ratcliff, W.C., Herron, M.D., Howell, K., Pentz, J.T., Rosenzweig, F., Travisano, M., 2013. Experimental evolution of an alternating uni- and multicellular life cycle in *Chlamydomonas reinhardtii*. *Nat. Commun.* 4, 2742.
- Rossetti, V., Schirmer, B.E., Bernasconi, M.V., Bagheri, H.C., 2010. The evolutionary path to terminal differentiation and division of labor in cyanobacteria. *J. Theor. Biol.* 262, 23–34.
- Rossetti, V., Filippini, M., Svercel, M., Barbour, A.D., Bagheri, H.C., 2011. Emergent multicellular life cycles in filamentous bacteria owing to density-dependent population dynamics. *J. R. Soc. Interface* 8, 1772–1784.
- Sanchez, A., Gore, J., 2013. Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biol.* 11 (4), e1001547.
- Santos, F.C., Pacheco, J.M., 2005. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* 95, 098104.
- Stolz, O., 1885. *Vorlesungen ueber allgemeine Arithmetik: nach den Neueren Ansichten*, Teubners, Leipzig, pp. 173–175.
- Szabó, G., Hauert, C., 2002. Phase transitions and volunteering in spatial public goods games. *Phys. Rev. Lett.* 89, 118101.
- Szabó, G., Fath, G., 2007. Evolutionary games on graphs. *Phys. Rep.* 446, 97–216.
- Tarnita, C.E., Ohtsuki, H., Antal, T., Fu, F., Nowak, M.A., 2009. Strategy selection in structured populations. *J. Theor. Biol.* 259, 570–581.
- Tarnita, C.H., Taubes, C.H., Nowak, M.A., 2013. Evolutionary construction by staying together and coming together. *J. Theor. Biol.* 320, 10–22.
- van Baalen, M., Rand, D.A., 1998. The unit of selection in viscous populations and the evolution of altruism. *J. Theor. Biol.* 193 (4), 631–648.
- Willensdorfer, M., 2008. Organism size promotes the evolution of specialized cells in multicellular digital organisms. *J. Evol. Biol.* 21, 104–110.
- Wilson, E.O., 1971. *The Insect Societies*. Harvard University Press, Cambridge.