(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Journal of Theoretical Biology 320 (2013) 10–22

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Evolutionary construction by staying together and coming together

Corina E. Tarnita [a,b,*], Clifford H. Taubes [c], Martin A. Nowak [d,e]

[a] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540, United States
[b] Harvard Society of Fellows, Harvard University, Cambridge, MA 02138, USA
[c] Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
[d] Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA
[e] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

HIGHLIGHTS

- Mathematical theory that explores how evolution can be constructive.
- Two fundamental operations: 'staying together' (ST) and 'coming together' (CT).
- ST-individuals form larger units by not separating after reproduction.
- CT-independent individuals form aggregates.
- ST and CT are very different mechanisms.

ARTICLE INFO

Article history:
Received 22 May 2012
Received in revised form
7 November 2012
Accepted 19 November 2012
Available online 30 November 2012

Keywords:
Multicellularity
Evolutionary game theory
Complexity

ABSTRACT

The evolutionary trajectory of life on earth is one of increasing size and complexity. Yet the standard equations of evolutionary dynamics describe mutation and selection among similar organisms that compete on the same level of organization. Here we begin to outline a mathematical theory that might help to explore how evolution can be constructive, how natural selection can lead from lower to higher levels of organization. We distinguish two fundamental operations, which we call 'staying together' and 'coming together'. Staying together means that individuals form larger units by not separating after reproduction, while coming together means that independent individuals form aggregates. Staying together can lead to specialization and division of labor, but the developmental program must evolve in the basic unit. Coming together can be creative by combining units with different properties. Both operations have been identified in the context of multicellularity, but they have been treated very similarly. Here we point out that staying together and coming together can be found at every level of biological construction and moreover that they face different evolutionary problems. The distinction is particularly clear in the context of cooperation and defection. For staying together the stability of cooperation takes the form of a developmental error threshold, while coming together leads to evolutionary games and requires a mechanism for the evolution of cooperation. We use our models to discuss simple aspects of the evolution of protocells, eukarya, multi-cellularity and animal societies.

1. Introduction

Evolution is constructive. Over time evolution has led from simple to complicated forms, from prokaryotic to eukaryotic cells, from single cellular to multicellular organisms, from solitary insects to colonies, from animal groupings to human society (Bonner, 1988, 1998; Carroll, 2001; Knoll, 2011; Lynch and Conery, 2003; Maynard Smith and Szathmary, 1998). Yet most mathematical descriptions of evolution so far have focused on competition on the same level of organization (Crow and Kimura, 2009; Eigen and Schuster, 1977; Ewens, 2010; Fontana and Buss, 1994; Hofbauer and Sigmund, 1998; May, 2001; Nowak, 2006). We think that it will likely be profitable to introduce mathematical models that explicitly distinguish two basic operations underlying evolution's propensity for construction, these being the operations of staying together (ST) and coming together (CT). As a first step to this end, we present in the body of this paper some simple mathematical models for both the ST and CT operations. These two operations appear to have been involved

* Corresponding author at: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540, United States.
E-mail address: ctarnita@princeton.edu (C.E. Tarnita).

0022-5193/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtbi.2012.11.022

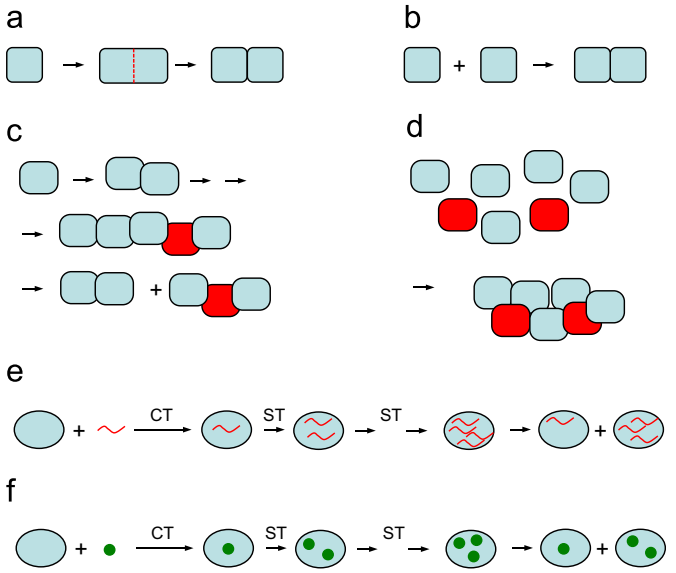

Fig. 1. (a, b) Staying together, ST, and coming together, CT, are fundamental operations which empower biological construction. (c, d) Many origins of multicellularity are based on ST, while others involve CT. (e) Protocells arise when lipid vesicles and RNA replicators come together. Subsequently the structure of the protocell enables the RNA sequences to stay together after replication. (f) The endosymbiotic theory is based on a coming together of different bacteria and a staying together of the endosymbionts after reproduction within the host.

in all great creative steps in the evolution of life on earth (Fig. 1). What follows directly in this introduction gives a handful of examples where the distinction between ST and CT might lead to novel insights.

The emergence of protocells requires the coming together of lipid membranes and informational polymers (Chen et al., 2004). The former generate replicating vesicles which provide spatially localized compartments. The latter give rise to replicating genomes that encode heritable information. The replication of polymers within vesicles leads to a staying together of parent and offspring polymers. Separation can occur during protocell division.

The endosymbiotic theory suggests that the emergence of the eukaryotic cell was caused by the coming together of different bacterial cells (Margulis, 1981). Endosymbiosis means that one cell makes a living inside another cell. The organelles of the eukaryotic cell, such as mitochondria and chloroplasts, are thought to have arisen in this way. The creative potential of CT is evident in this process, as the symbiotic fusion can involve cells with very different properties thereby leading to entirely new organisms. The reproduction of the endosymbiont within its host cell causes a staying together of the endosymbiont.

Multicellularity has evolved many times and in all three domains of life (archea, bacteria, eukaria) (Bell, 1997; Bonner, 2008; Boraas et al., 1998; Furusawa and Kaneko, 2000; Grosberg and Strathmann, 2007; Hall-Stoodley et al., 2004; King, 2004; Kirk, 2003, 2005; Kolter, 2010; Michod, 2007; Michod and Roze, 2001; Pfeiffer and Bonhoeffer, 2003; Rainey and Kerr, 2010; Rossetti et al., 2010; Stanley, 1973; Webb et al., 2003; Willensdorfer, 2008; Wolpert, 1990). The evolution of multicellularity from unicellular ancestors led to macroscopic forms with new body plans, higher grades of morphological complexity and was often followed by periods of rapid diversification. It has been suggested that aquatic organisms often evolved multicellularity as the products of cell division failed to separate (ST), while several terrestrial origins involved motile aggregation of cells or nuclei (CT) (Bonner, 1998). Multicellularity allows the subsequent evolution of cellular differentiation and division of labor. In cyanobacteria, for example, specialized heterocysts occur at regular intervals within filaments and perform nitrogen fixation, while the surrounding vegetative cells engage in photosynthesis.

Another major constructive event in the evolutionary unfolding of life on earth is the emergence of animal societies (Alexander, 1974; Krebs and Davies, 1991; Gadagkar, 2001; Gadagkar and Bonner, 1994; Hölldobler and Wilson, 2009; Hunt, 2007; Leadbeater et al., 2011). The two mechanisms that we describe in this paper are identified as the two primary routes by which animal societies form (Krebs and Davies, 1991). Coming together occurs when individuals aggregate because of inherent advantages of group living—for example, increased alertness and defense against predators (as seen in groups of baboons) and increased capabilities for detecting and harvesting difficult to locate food resources (as seen in lions, wolves or wild dogs) (Alexander, 1974). Staying together occurs when offspring delay dispersal and remain with their natal group because of various constraining factors that restrict their option of dispersing and breeding on their own. This is the case in many cooperatively breeding birds and mammals where offspring delay dispersal and help at the nest of their parents. Delayed dispersal can happen, for example, because offspring are unable to find nesting sites of their own (Alexander, 1974; Leadbeater et al., 2011). ST might arise for instance when mutations or behavioral modifications prevent offspring from leaving the nest. This is the more likely route for the origin of eusociality in insects. There are notable similarities when comparing the evolution of social insects and multicellularity; in both cases we find large variation in group size ranging from solitary reproduction to complex colonies with division of labor (Gadagkar and Bonner, 1994). While ST is the more likely route to eusociality, CT occurs in certain cases where multiple fertilized females join forces to establish a new colony (Gadagkar, 2001; Hölldobler and Wilson, 2009; Hunt, 2007). ST is the step that leads from subsocial to eusocial (Nowak et al., 2010; Hunt, 2011), while CT can lead from parasocial to primitively eusocial (Gadagkar, 2001).

In many examples of social aggregations both CT and ST operate at the same time; for example, the offspring of cooperatively breeding birds stay with the nest, thus performing ST, but females immigrate and join new nests which is a form of CT. Similarly, prides of lions are formed of sisters (ST) joined by immigrant males (CT). While often both mechanisms are employed simultaneously, a study of each mechanism independently can provide important insight into the similarities and differences between them. Based on this understanding, more complex scenarios can then be analyzed. To exemplify the way our framework can be used to describe situations where both staying together and coming together occur simultaneously we use an existing classification of social behaviors in insects (Michener, 1974) and give a parallel classification in terms of ST and CT (Fig. 2).

In this paper we are concerned with the primitive forms of construction. Once construction has been achieved, one can discuss the maintenance of complexes, the evolution of specialization, the point of no return (beyond which a reversal to a solitary state is impossible due to complete specialization) (Crespi, 2008) and the progression to higher dimensional phenotype spaces, which can promote the evolution of biological diversity (Doebeli and Ispolatov, 2010). While this framework can still be employed to address some of these issues, these are not the questions we are trying to address here. What we are concerned with in this paper is the very first stage of construction and how that might occur and be selected for.

This paper proposes a framework in which biological construction can be analyzed, but we do not study a specific model system. In many ways our models are only very preliminary steps

Level of Sociality	Cooperative Brood Care	Overlapping generations	Division of labor	ST/CT framework
Solitary	-	-	-	solitary
Subsocial	-	-	-	preST
Communal	-	-	-	preCT
Quasisocial	+	-	-	CT+preST
Semisocial	+	-	+	CT + preST + division of labor
Eusocial (subsocial route)	+	+	+	preST ⟶ ST + division of labor
Eusocial (parasocial route)	+	+	+	CT+preST ⟶ CT+ST + division of labor

Fig. 2. A classification using the CT/ST framework of the different types of social behavior encountered in insects. We define preST to be a precursor state of staying together (ST). In the case of social insects, this corresponds to subsocial species: a mother feeding her daughters until they are grown, after which both the mother and the young disperse. We define preCT to be a precursor state of coming together (CT). This denotes an aggregation whose units do not interact, despite being in close proximity. For social insects, this corresponds to communal breeders. Then quasisocial species which are aggregations of adult mothers engaging in cooperative brood care, followed by dispersal of the young when reaching maturity can be described as a CT event followed by a preST event. Semisocial behavior is the same as quasisocial, except that it also has division of labor. Eusocial behavior developed via the subsocial route represents a transition from preST to ST. Eusocial behavior developed via the parasocial route represents a transition from CT+preST to CT+ST. The latter seems to be rarely found and is usually referred to as primitive eusociality. In both cases, eusociality also has division of labor.

to introduce ST and CT for haploid systems and to show that a study of the differences between the two processes is worthwhile. We do not claim that our models can be immediately applied to the examples discussed above but we hope that our framework can be extended and eventually applied to study specific model systems. Any such application will likely require adapting our framework to the particular system and adding layers of specificity to the bare backbone we provide. However, in doing so, one could make predictions, use the framework to analyze them theoretically, set up experiments to test empirically and then compare theory and practice.

Maynard Smith and Szathmary (1998) offer qualitative discussions of these mechanisms in the context of major transitions in evolution; see also Queller (1997, 2000). Grosberg and Strathmann (2007) provide an extensive review of mechanisms for the evolution of multicellularity. Theoretical discussions for the evolution of multi-cellularity are given by Michod and Roze (1999) and Michod (2007). These discussions present the evolution of multi-cellularity either as a kin selection or as a multilevel selection problem where groups of cooperating individuals become so integrated that they evolve into new higher-level individuals (Michod, 2007). Both ST and CT are mentioned, but are treated within the same approach. In this paper we show that there is a significant difference between staying together and coming together and their respective outcomes.

The paper is organized as follows. In Section 2 we discuss the mathematical model for staying together. In Section 3 we discuss the model for coming together. In Section 4 we compare the mathematical insights obtained in the previous two sections for ST and CT. In Section 5 we address the problem of cooperation and conflict for both ST and CT. One of the advantages of complex formation is the potential for the exploration of new niches. In Section 6 we explore how inhabiting a new niche can facilitate the evolution of construction. We conclude by discussing in Section 7 the importance of the ST–CT framework and the possible ways in which it can be employed to illuminate the idea of construction in biology.

2. Staying together (ST)

In this section we formulate a theory for staying together (ST). There are two types of units, A and B. Following reproduction, the former can make complexes of increasing size, while the latter always separate. For simplicity, we assume that q is independent of the size of the complex. Let A_i denote complexes of size i; they produce new units at rate a_i. B reproduces at rate b. Complexes that result from staying together can have different ways of reproduction. One way of reproduction, which we will simply call ST, behaves such that an A_i complex produces a new cell A which then can start forming its own complex.

A second way of reproduction exists which can be seen for instance in chain forming bacteria (e.g. cyanobacteria). There, a filament fissions into two filaments of arbitrary size (this is a generalization of the above case; occasionally a single cell might break away, but bigger aggregates can also break away). We call this process "ST with chain breaking".

Finally we identify a third way of reproduction, which we call "ST protocells". This is inspired by the emergence of protocells. The emergence of protocells requires the coming together of lipid membranes and informational polymers. The former generate replicating vesicles which provide spatially localized compartments. The latter give rise to replicating genomes that encode heritable information. To describe the formation of protocells in terms of ST, let A be an RNA inside a lipid vesicle. Let B be free RNA (outside any vesicles). We will denote by A_i a vesicle that contains i copies of RNA. Inside the vesicle, the RNA makes more copies of itself. At rate r, the lipid vesicle can divide into two vesicles, each containing some of the RNAs from the larger vesicle.

Below we discuss these three possibilities.

2.1. ST

In this section we look at one possibility for how ST could work. There are two types of units, A and B. Following reproduction, the former stay together with probability q, while the latter always separate. Let A_i denote aggregates of size i; they produce new units at rate a_i and die at rate d_i. B reproduces at rate b and dies at rate d. The parameters a_i and b constitute the fitness landscape. Different fitness landscapes can lead to different outcomes. We have the following system of 'biological reactions':

$$A_i \xrightarrow{a_i q} A_{i+1}$$

$$A_i \xrightarrow{a_i(1-q)} A_i + A$$

$$A_i \xrightarrow{d_i} 0$$

$$B \xrightarrow{b} B+B$$

$$B \xrightarrow{d} 0 \tag{1}$$

These reactions can be captured by the following system of differential equations, where x_i denotes the abundance of complex A_i and y denotes the abundance of B:

$$\dot{x}_1 = (1-q)\sum_i a_i x_i - q a_1 x_1 - d_1 x_1 - \phi x_1$$

$$\dot{x}_i = q(a_{i-1}x_{i-1} - a_i x_i) - d_i x_i - \phi x_i \quad i = 2,3,\ldots$$

$$\dot{y} = (b-d)y - \phi y \tag{2}$$

The additional death term ϕ is chosen such that the total abundance of units (not of complexes) is constant: $y+\sum_i i x_i = \text{constant}$. Without loss of generality we will assume this constant to be 1. Then the ϕ that fulfills the condition $y+\sum_i i x_i = 1$ is $\phi = (b-d)y + \sum_i (a_i - i d_i)x_i$. For the system studied here we cannot give a general explicit formula for ϕ, but only an implicit one as will be seen below.

Because we assume that the two types compete for the same niche, this system is characterized by competitive exclusion (for a generic choice of parameters). The system has two equilibria—an all-x equilibrium and an all-y equilibrium. Depending on the parameters, one of them will be stable and one unstable with respect to invasion by the other. This means that either A or B wins while the other is driven to extinction. In order to analyze this system, we calculate ϕ for the all-x equilibrium and for the all-y equilibrium. The equilibrium that has the larger ϕ is stable.

At the all-y equilibrium we find $\phi = b-d$. At the all-x equilibrium we find ϕ to be given by

$$1 = \frac{1-q}{q} \sum_{n\geq 1} \prod_{k=1}^{n} \frac{q a_k}{\phi + d_k + q a_k} \tag{3}$$

Then the condition that A wins is equivalent to $\phi > b-d$ which gives

$$\frac{1-q}{q} \sum_{n\geq 1} \prod_{k=1}^{n} \frac{q a_k}{b-d+d_k+q a_k} \geq 1 \tag{4}$$

Further analysis is not possible without deciding on a "death landscape" (i.e. how d_i depends on i) and this needs to be specified for the particular system under study. Throughout the rest of the paper we will study the case $d_i = d = 0$ because the conclusions that we are interested in (e.g. competitive exclusion) are preserved in this case. Moreover, this choice allows us to focus on the fitness landscape. In this case, the equations become

$$\dot{x}_1 = (1-q)\sum_i a_i x_i - q a_1 x_1 - \phi x_1$$

$$\dot{x}_i = q(a_{i-1}x_{i-1} - a_i x_i) - \phi x_i, \quad i = 2,3,\ldots$$

$$\dot{y} = by - \phi y \tag{5}$$

The death rate ϕ is chosen, as before, such that the total abundance of units is constant: $y+\sum_i i x_i = 1$. It is easy to see that the ϕ that fulfills this condition is $\phi = by + \sum_i a_i x_i$.

The use of ϕ is exactly as in the quasispecies equation and in the replicator equation (Eigen and Schuster, 1977; Hofbauer and Sigmund, 1998). We note that all A complexes have exactly the same death rate, ϕ, independent of their size; a singleton B has the same death rate, ϕ as well. This corresponds to a system in which complexes disappear because of dilution, not because of actual death—for example, a small pond in which water continually flows in and out or a flow reactor in a lab. For a discussion of other possible types of density limitation see Appendix A.

Substituting $d_i = d = 0$ in the condition above, we find the condition that A wins

$$\frac{1-q}{q} \sum_{n\geq 1} \prod_{k=1}^{n} \frac{q a_k}{b + q a_k} \geq 1 \tag{6}$$

In the special case where $a_1 = b = 1$ and $a_k = a$ for all $k \geq 2$ we find that A grows faster than B if $a > 2/(1-q)$.

Another interesting special case is $a_k = kb$ for all k. Here the two strategies are neutral in the sense that they both have the same ϕ value, $\phi = b$. To prove neutrality, note that we chose ϕ such that $y + \sum i x_i = 1$. Then $\phi = by + \sum a_i x_i = b(y + \sum i x_i) = b$. One immediate implication of this result is that in order for A to outperform B, at least one a_i has to be greater than b. In other words, at least one of the A_i aggregates has to be better at producing As than i copies of the single B are at producing Bs. Note that for the case when the death rates are distinct for complexes of different sizes the neutrality condition is $a_i - i d_i = i(b-d)$ for all i.

If $a_1 < b$ then there is a cost for having the ability to do ST when comparing the reproduction rate of single units. In this case, one can find landscapes where ST is not "fast enough" in forming aggregates and hence A loses against B regardless of what q is. It is in general very hard to classify all landscapes for which ST is not "fast enough"; however, we can study a sufficiently general case that will exemplify our claim. Consider the landscape $a_1 = \cdots = a_k = \alpha_1 b$ and $a_i = \alpha_2 b$ for all $i \geq k+1$, such that $\alpha_1 < 1 < \alpha_2$. Then condition (6) becomes

$$\alpha_1 + (\alpha_2 - \alpha_1)\left(\frac{q\alpha_1}{1+q\alpha_1}\right)^k > \frac{1}{1-q} \tag{7}$$

Since $q < 1$ and $\alpha_1 < 1$ then $q\alpha_1/(1+q\alpha_1) < 1/2$ which means that the left hand side of the above condition is less than $(\alpha_1+\alpha_2)/2$. It is then easy to choose $\alpha_1 < 1 < \alpha_2$ such that $(\alpha_1+\alpha_2)/2 < 1$ and since $1/(1-q) > 1$ for all q, condition (7) will never be fulfilled, for any q. For simplicity we made a very course argument above, but in fact for a very large range of parameters $\alpha_1 < 1 < \alpha_2$, condition (7) will not hold for any q.

If however $a_1 > b$ then even as a single unit, A is better than B at making more copies of itself. Hence it wins even for $q = 0$ (this means that there is no need for A to make aggregates because even in the single unit phase it is nevertheless better than B). Thus, the most interesting case to study is $a_1 \leq b$.

Finally, another interesting fact about ST is that the average size of the A aggregates at equilibrium does not depend on a_i and is equal to $1/(1-q)$. The average size of the system is given by $(\sum_i i x_i)/(\sum_i x_i) = 1/(\sum_i x_i)$, since $\sum_i i x_i = 1$. To prove this statement we use the following two equilibrium equations:

$$(1-q)\sum_{i\geq 1} a_i x_i - (q a_1 + \phi)x_1 = 0 \tag{8}$$

$$q a_{i-1} x_{i-1} - (q a_i + \phi) x_i = 0 \tag{9}$$

Summing (9) for $i \geq 2$ and using the fact that $\phi = \sum a_i x_i$ we find that $\sum_{i\geq 2}(q a_i + \phi)x_i = q\phi$ and therefore $q\phi - (q a_1 + \phi)x_1 + \phi \sum_{i\geq 1} x_i = q\phi$. This together with (8) yields

$$\phi \sum_{i\geq 1} x_i = (q a_1 + \phi)x_1 = (1-q)\phi \tag{10}$$

Hence, the average size of the system becomes $\sum i x_i / \sum x_i = 1/(1-q)$. In the more general case when we allow for different death rates, the corresponding finding is that $\sum i(d_i+\phi)x_i / \sum (d_i+\phi)x_i = 1/(1-q)$.

2.2. ST with chain breaking

In this section we analyze staying together with chain breaking. Again, there are two types of units, A and B. Let a_i be the rate at which an aggregate of size i turns into an aggregate of size $i+1$. Let c be the rate constant at which an aggregate breaks. The breaking is random—a string of size i can break at $i-1$ many places. We have the following 'biological reactions':

$$A_i \xrightarrow{a_i} A_{i+1}$$
$$A_i \xrightarrow{c} A_k + A_{i-k} \quad i > k \geq 1$$
$$B \xrightarrow{b} B + B \tag{11}$$

These reactions can be represented by the following system of differential equations, where x_i is the frequency of aggregate A_i.

$$\dot{x}_1 = -a_1 x_1 + 2c \sum_{j=2}^{\infty} x_j - \phi x_1$$
$$\dot{x}_i = a_{i-1} x_{i-1} - a_i x_i - c(i-1) x_i + 2c \sum_{j=i+1}^{\infty} x_j - \phi x_i \quad i = 2,3,\ldots$$
$$\dot{y} = by - \phi y \tag{12}$$

As before, ϕ is chosen such that $y + \sum_i i x_i = 1$. Hence $\phi = by + \sum_i a_i x_i$. For $a_i = ib$ we can show neutrality as before. Using the first equation and summing up the second equation at equilibrium for $i \geq 2$ we find that the average size of the system $\sum_i i x_i / \sum x_i = 1 + \phi/c$. This gives a very simple and interesting relationship between the average size and ϕ. As for the ST model, we find competitive exclusion. The evolutionary outcome depends on the value of c, the rate of chain breaking. We will exemplify this statement by looking at a landscape like the one in Fig. 3: in such a landscape, the ability to form chains has a cost which is overcome by chains of intermediate length; long chains are inefficient. In this case, for very low or very high values of c, ST is outcompeted. If the rate of chain breaking is too high, then the complexes break too often and they do not get the benefit of reaching the intermediate size; if, on the other hand, the rate of chain breaking is too low, then complexes become too large which, by assumption, is inefficient.

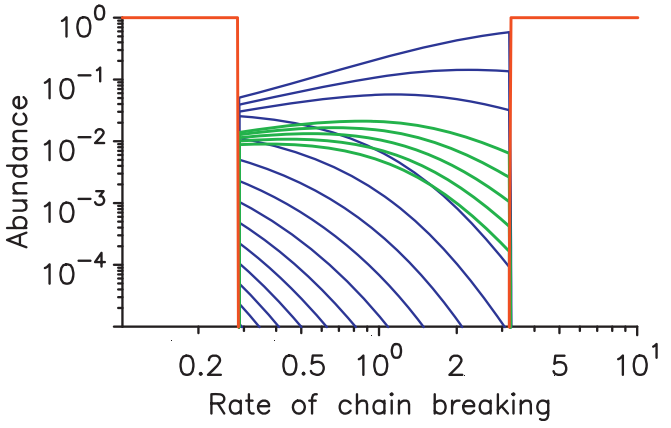

Fig. 3. A model of staying together, which is inspired by simple organisms that form multi-cellular chains. Cell division increases the length of a chain. Chains can break into two pieces thereby increasing the number of chains. A chain of length n can break in $n-1$ places. The rate of chain breaking is the same for each place and given by c. The figure shows the equilibrium configuration. As for the standard ST model there is competitive exclusion: A and B cannot coexist. We choose a fitness landscape where the ability to form chains has a cost, which is overcome by chains of intermediate length. Longer chains are inefficient. In particular we assume: $a_i = 0.95i$ for $i = 1,2,3$; $a_i = 2i$ for $i = 4,..,8$ and $a_i = 0.3i$ for all $i \geq 9$. The evolutionary outcome depends on the rate of chain breaking, c. A wins if approximately $0.29 < c < 3.2$. Otherwise B wins. If c is too small then long, inefficient chains accumulate. If c is too large, then chains of optimum length ($i = 4,\ldots,8$) are rarely formed. Color code: green for x_i with i=4..8; blue for all other x_i (1, 2, 3, 9, 10,…from top to bottom); red for y; equilibrium abundances are shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Other important theoretical and experimental work which addresses ST with chain break in filamentous bacteria is that done by Rossetti et al. (in press). The way they model ST with chain breaking is similar to our own approach. However, there are some differences in the setup. Rossetti et al. (in press) assume that every cell has the same birth rate regardless of the length of the chain, but dependent on the total population size; moreover, complexes can only break when cells die and the death rate of a cell is again independent of the size of the complex it belongs to, but dependent on the total population size. There are also differences in the questions we ask: while Rossetti et al. (in press) want to understand distributions of filament lengths without a priori assuming any benefits of complexes compared to multicellular life, we are more concerned with the conditions under which filaments can outcompete their solitary ancestors. While Rossetti et al. (in press) often find coexistence at equilibrium, we look at situations where competitive exclusion is present. Both questions are very relevant to the study of the emergence of multicellularity.

2.3. ST protocells

The emergence of protocells requires the coming together of lipid membranes and informational polymers. The former generate replicating vesicles which provide spatially localized compartments. The latter give rise to replicating genomes that encode heritable information. Below we describe the formation of protocells in terms of ST.

Let A be an RNA inside a lipid vesicle. Let B be free RNA (outside any vesicles). We will denote by A_i a vesicle that contains i copies of RNA. Inside the vesicle, the RNA makes more copies of itself. At rate r, the lipid vesicle can divide into two vesicles, each containing some of the RNAs from the larger vesicle. We can then write the following biological reactions:

$$A_i \xrightarrow{a_i} A_{i+1}$$
$$A_i \xrightarrow{r_{ik}} A_k + A_{i-k} \quad i > k \geq 1$$
$$B \xrightarrow{b} 2B \tag{13}$$

where

$$r_{ik} = r \binom{i}{k} \frac{1}{2^i} \tag{14}$$

Although the process seems similar to that of ST with chain breaking, the dynamics and consequently the equations describing the dynamics are different. This comes from the fact that a chain of length i can break into a chain of size k and one of size $i-k$ in two ways whereas it is natural to assume that the RNA molecules inside the cell are randomly distributed over the daughter cells which means that the breaking of a cell with i molecules into two cells with $i-k$ and k molecules respectively, can happen in $\binom{i}{k}$ ways. We thus obtain the following system of equations:

$$\dot{x}_1 = -a_1 x_1 + r \sum_{k \geq 1} \frac{k}{2^{k-1}} x_k - r x_1 - \phi x_1$$
$$\dot{x}_i = a_{i-1} x_{i-1} - a_i x_i + r \sum_{j \geq i} \frac{1}{2^{j-1}} \binom{j}{i} x_j - r x_i - \phi x_i$$
$$\dot{y} = by - \phi y \tag{15}$$

As usual, the density dependent death rate is chosen such that $y + \sum i x_i = 1$ which yields $\phi = by + \sum a_i x_i$. Again, it is

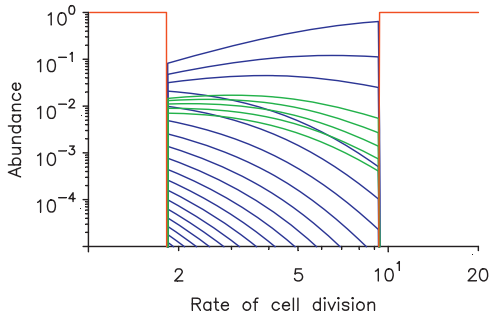

Fig. 4. A model of staying together, which is inspired by protocells. Here i denotes the number of RNA molecules inside a cell. This number increases by RNA replication. When a cell divides, the RNA molecules are randomly distributed over the two daughter cells. This cellular RNA reproduction is in competition with solitary RNAs that reproduce outside of cells. The evolutionary outcome depends on the fitness landscape and the rate of cell division, r. We choose a fitness landscape where protocells that harbour an intermediate number of RNAs have a maximum replication rate. If the number of RNAs in a protocell is too small then their catalytic activity (for example to bring nucleotides into the protocell) is too low. If the number of RNAs in a protocell is too high then there is too much competition for these nucleotides. As a simple numerical example we assume: $a_i=0.95i$ for $i=1,2,3$; $a_i=2i$ for $i=4,..,8$ and $a_i=0.3i$ for all $i\geq 9$. The replication rate of RNA outside of cells is 1. There is competitive exclusion between A (RNA inside protocells) and B (RNA outside protocells). In our example A wins if approximately $1.8<r<9.4$. If r is too low then the amount of RNA inside protocells is too large. If r is too high then the average amount or RNA inside protocells is below maximum efficiency. Color code: green for x_i with $i=4..8$; blue for all other x_i (1, 2, 3, 9, 10,…from top to bottom); red for y; equilibrium abundances are shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

straightforward to see that if $a_i=bi$ for all $i\geq 1$, the system is neutral, since $\phi=b$.

In Fig. 4 we perform simulations to show how the evolutionary outcome of the competition between free RNAs and RNAs contained in vesicles depends on the rate r of cell division. For comparison, we have chosen to look at the same fitness landscape as for ST with chain breaking. As before, we find competitive exclusion. For small and large rates of cell division the solitary RNAs win: for small rates this is because the protocells grow too big and inefficient; for large rates of cell division it is because the protocells divide too much and they are too small, where again it is costly and inefficient. Thus for this fitness landscape protocells win only for intermediate rates of cell division.

2.4. Example: ST with maximum size $n=2$

To conclude the discussion of ST, let us consider the following simple example of ST, where complexes can have at most size 2. Then the reactions can be written as

$$\begin{aligned} &A\xrightarrow{a_1 q} AA\\ &A\xrightarrow{a_1(1-q)} A+A\\ &AA\xrightarrow{a_2} AA+A\\ &B\xrightarrow{b} 2B \end{aligned} \tag{16}$$

Note that in this case however, we cannot simply use the general result derived above to find the condition for ST to win; this is because we impose the cutoff $n=2$ and hence we need to write a different system of equations to describe the dynamics:

$$\begin{aligned} \dot{x}_1&=(1-2q)a_1x_1+a_2x_2-\phi x_1\\ \dot{x}_2&=qa_1x_1-\phi x_2\\ \dot{y}&=by-\phi y \end{aligned} \tag{17}$$

The condition for the x equilibrium to be stable under invasion is

$$b^2-ba_1(1-2q)-qa_1a_2<0 \tag{18}$$

If moreover $b=a_1$ then the above condition becomes $a_2>2a_1$. For the system with the ϕ density limitation we can, without loss of generality, take $b=1$. Then the condition that the x equilibrium is stable is

$$q>\frac{1-a_1}{a_1(a_2-2)} \tag{19}$$

Hence the success of ST depends on the probability of staying together being large enough. If this is the case, then construction by ST will outcompete solitary life. Since q is a probability it is also bounded upwards by 1; this leads to the condition that $a_2>1+1/a_1$. Thus, for ST there are a_i values where no q would exist that is large enough for construction to be favored.

3. Coming together (CT)

Next we formulate a theory for coming together (CT). Again there are two types of units. A units can combine with each other at rate β to form complexes, while B units are always solitary. We mentioned in the introduction that one of the most interesting aspects of CT is that different types can come together to form heterogenous complexes. In this paper we do not explore how heterogeneous complexes form and behave, except in the context of cooperation and defection where we look at what happens when a defector type can join the complexes. Otherwise we only look at homogenous complexes formed with one type. In this case, the only difference between ST and CT is that CT requires the presence of other individuals of the same type whereas ST grows from one founder individual. As before, A_i complexes produce new units at rate a_i, while B reproduces at rate b. We obtain

$$\begin{aligned} &A_i+A\xrightarrow{\beta} A_{i+1}\\ &A_i\xrightarrow{a_i} A_i+A\\ &B\xrightarrow{b} B+B \end{aligned} \tag{20}$$

These reactions can be described by the following system of differential equations, where x_i is the abundance of aggregate A_i and y is the abundance of B:

$$\begin{aligned} \dot{x}_1&=\sum_i a_ix_i-\beta x_1\sum_i x_i-\beta x_1^2-\phi x_1\\ \dot{x}_i&=\beta x_1(x_{i-1}-x_i)-\phi x_i \quad i=2,3,\ldots\\ \dot{y}&=by-\phi y \end{aligned} \tag{21}$$

As before, the density limitation is chosen such that $y+\sum ix_i=1$. We then find that $\phi=by+\sum a_ix_i$. As for ST we find that the fitness landscape $a_i=ib$ for all i is neutral since $\phi=by+\sum a_ix_i=b(y+\sum ix_i)$.

Unlike in the ST system where we find competitive exclusion, in the CT system there can exist mixed equilibria or multiple equilibria (see Fig. 5). While in general finding all these equilibria is a hard problem, we can show that the all-x equilibrium is unique. If $y=0$ and ϕ is such that $\sum_{i\geq 1}ix_i=1$, then we must have that $\phi=\sum_{i\geq 1}a_ix_i$. Moreover, from the second equation we find that $x_i=(\beta x_1/(\beta x_1+\phi))^{i-1}x_1$. Hence, we can rewrite the first equation as $\phi^2(1-x_1)-2\phi\beta x_1^2-\beta^2x_1^3=0$. We can solve this for ϕ to find that it must have a unique positive root

$$\phi=\beta x_1\frac{\sqrt{x_1}}{1-\sqrt{x_1}} \tag{22}$$

Moreover, we know that

$$\phi=\sum a_ix_i=x_1\sum a_i\left(\frac{\beta x_1}{\beta x_1+\phi}\right)^{i-1}=x_1\sum a_i(1-\sqrt{x_1})^{i-1} \tag{23}$$

The last equality comes from substituting ϕ in the denominator by the expression in (22). Now denoting $1-\sqrt{x_1}=w$ and equating

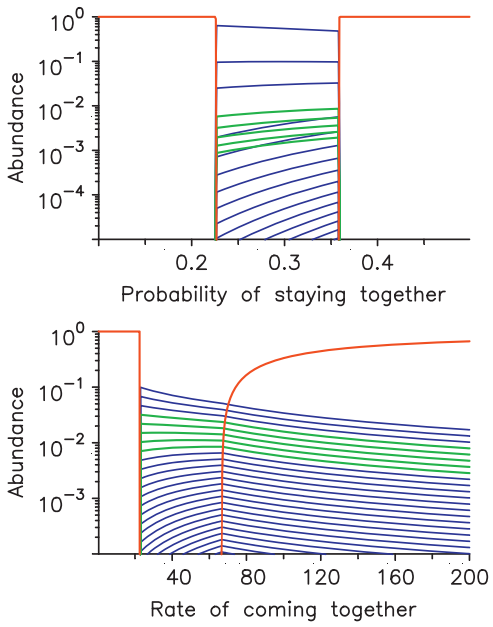

Fig. 5. Natural selection of construction. We assume that complexes of intermediate size have a fitness advantage. Complexes that are too small or too large have a lower per capita reproductive rate than singletons. The curves represent the equilibrium abundances of the x_i and y, denoting respectively A and B. (a) For ST we show equilibrium abundances as function of q, the probability of staying together. For small and large values of q we observe that B wins because complexes are either too small or too large. For intermediate values of q we observe that A wins. The ST model used is the one discussed in Section 2.1 with parameters $a_i=0.95i$ for $i=1,2,3$; $a_i=2i$ for $i=4,..,8$ and $a_i=0.3i$ for all $i\geq 9$. (b) For CT we show equilibrium abundances as a function of the rate of coming together, β. If β is too small then B is selected. For intermediate values of β we observe that A is selected. For larger values of β we observe a mixed equilibrium. Note that the all-B equilibrium is always stable (not shown). Parameters: $a_i=1.7i$ for $i=4,..,8$; and $a_i=0.5i$ for all other i. In both cases we assume that only complexes of intermediate size have a reproductive rate (per unit) which is faster than that of singletons, $b=1$. This assumption represents the idea that there is a cost of complex formation which is only overcome above a certain size, but large complexes are again inefficient. Color code: green for x_i with $i=4..8$ (from top to bottom); blue for all other x_i ($i=1,2,3,9,10,\ldots$ from top to bottom); red for y; equilibrium abundances are shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

the two expressions for ϕ we find that the equilibrium x_1 has to be a zero of the function $g(w)=-1+(a_1+\beta)w+a_2w^2+a_3w^3+\cdots$. Since all $a_i>0$ and $\beta>0$, then g is an increasing function on the interval $[0,\infty)$ and thus if it has a solution, it must be unique. Since $g(0)=-1<0$ and $g(1)>0$, a solution must exist on the interval $[0,1]$. Hence the equilibrium is unique.

In general it is hard to determine which of the equilibria are stable. The only case we can easily analyze is for the all-y equilibrium. To determine its stability, we set $\phi=b-d$ and determine whether an ϵ amount of x_1 can grow. We find that the all-y equilibrium is stable if and only if $a_1<b$.

3.1. Example: CT with maximum size $n=2$

As for ST, we now analyze the case where complexes can only reach size $n=2$.

$$A+A\xrightarrow{\beta}AA$$
$$A\xrightarrow{a_1}A+A$$
$$AA\xrightarrow{a_2}AA+A \tag{24}$$

Also, as in the case of ST, note that we cannot simply use the general result derived above to find the condition for CT to win; this is because we impose the cutoff $n=2$ and hence we need to write a different system of equations to describe the dynamics

$$\dot{x}_1=a_1x_1+a_2x_2-2\beta x_1^2-\phi x_1$$
$$\dot{x}_2=\beta x_1^2-\phi x_2$$
$$\dot{y}=y(b-d) \tag{25}$$

The $y=0$ equilibrium is given by

$$x_1=\frac{a_1+\sqrt{a_1^2+4a_2\beta}}{4\beta+a_1+\sqrt{a_1^2+4a_2\beta}}$$
$$x_2=\frac{2\beta}{4\beta+a_1+\sqrt{a_1^2+4a_2\beta}} \tag{26}$$

The condition that the $y=0$ equilibrium is stable is $\phi>b$

$$\frac{(a_1+\sqrt{a_1^2+4a_2\beta})^2}{2(4\beta+a_1+\sqrt{a_1^2+4a_2\beta})}>b \tag{27}$$

If $b=a_1$ this becomes $a_2>2a_1$. If $b=1$ then the most realistic scenario that we are interested to explore is $a_1<1$ and $a_2>2$. In this case, the $y=0$ equilibrium is stable if

$$\beta>\frac{(1-a_1)(a_2-2a_1)}{(a_2-2)^2} \tag{28}$$

Thus the success of CT depends on the ability β to find individuals with which one can form aggregates. Note that unlike in the ST case, in the CT case there is always a β that is large enough for construction to prevail, provided that $a_2>2$.

4. Discussion of ST versus CT

Although our treatment of CT in this paper is preliminary, in principle CT can allow the aggregation of different types (either individuals of the same species employing different strategies, or even individuals of different species as is for example the case in endosymbiosis). Thus unlike ST which is a clonal process (in the absence of mutation or sex), CT allows for the formation of complexes with varying degrees of relatedness. When the aggregates are formed by different types coming together, the CT equations can be interpreted as describing evolutionary game dynamics (Hofbauer and Sigmund, 1998) of multi-person games. The units in a complex are engaged in an interaction which generates payoff. The population contains game playing groups of different sizes. Solitary individuals represent loners. Thus, CT leads into a complexity comparable to all of evolutionary game theory; we believe this has to be explored in many subsequent studies.

Our mathematical approach differs from standard coagulation–fragmentation theory (CFT), which studies the formation and breaking-up of clusters formed by individual molecules, cells or animals (Gueron, 1998; Gueron and Levin, 1995; Gueron et al., 1996; Okubo, 1986). In contrast to CFT, we consider reproduction of individual units. The offspring can either stay with their cluster (staying together) or join other clusters (coming together). Furthermore, we study competition (natural selection) between reproductive strategies that have the ability to form clusters and those that do not. It would be of interest to explore extended CFT models that contain these additional operations.

ST leads to linear selection equations, while CT generates nonlinear equations which resemble aspects of evolutionary game dynamics. Once construction has been achieved by either one or both of those mechanisms, complex behavior like specialization of different units will follow. Our framework can be employed to study these subsequent steps, but here we focus only on the origin of construction and how simple complexes can

outcompete solitary life. Hence the crucial question that we want to address is: under which conditions is A favored over B by natural selection? The answer depends on the reproductive rates, a_i, which are an indicator of how much better or worse a complex of size i is at producing new units compared to a solitary individual, and on the parameters q and β which are the respective rates of complex formation.

In this paper we do not make any assumptions about whether the first constructive steps are beneficial or not. In fact, the more interesting scenarios arise when one allows the first constructive steps to be neutral ($a_i = ib$) or even detrimental ($a_i < ib$). However, ultimately only if certain steps are actually beneficial (i.e. $a_i < ib$ for small i but $a_i > ib$ for larger i), can natural selection promote A over B and thereby favor the evolution of construction.

Since we so far studied only the case where competition is for the same niche, for ST models we find competitive exclusion: stable coexistence between A and B is not possible. This means that for generic parameter values the system either converges to all-A or all-B (Fig. 5a). Without the density limitation, the ST system is a simple linear system with exponential growth. Thus, intuitively coexistence is not possible. This property is preserved even when we add the density limitation ϕ. For CT we find that multiple and mixed equilibria are possible. There can be stable coexistence between A and B or competitive exclusion (Fig. 5b).

A non-generic situation is given by the reproductive rates $a_i = ib$; in this case we find neutrality between A and B for both ST and CT. This result is intuitive: if the individual units in all complexes reproduce at the same rate as single units, then complex formation has neither an advantage nor a disadvantage. The emergence of complexes is then a consequence of neutral drift. For the origin of multicellularity Bonner has argued that the first step toward complex formation could have been neutral drift, while subsequent mutations lead to cellular differentiation, division of labor and discovery of new ecological niches (Bonner, 1998).

If there is a cost for the ability to form complexes in the sense that a single A unit reproduces at a slower rate than a single B unit, $a_1 < b$, then the all-B equilibrium is always stable under CT. This is because CT requires the presence of others in order to form complexes. Hence A can only evolve if an invasion barrier is overcome: a rare A mutant must locally drift to a minimum abundance before large enough complexes can arise which realize any possible selective advantage of CT. The A units must have the possibility to find each other. Note that a stochastic model would be needed to describe the drift to overcome the invasion barrier. The deterministic model that we look at only shows that there is an invasion barrier: the abundance of A can only increase if it starts above a critical threshold. In contrast, for ST the all-B equilibrium can be unstable even if $a_1 < 1$, which means that a rare A mutant can be immediately favored by natural selection. This is because each complex is formed by clonal reproduction starting from a single individual.

In spite of this apparent ease of construction using ST, there are many choices for the a_i parameters that allow evolution of construction for CT and not for ST. If, for instance, only large complexes provide fitness advantages, then high β values can facilitate the formation of such complexes via CT, but even the limit $q \to 1$ could be insufficient for ST. The reason is that ST can be too slow to form large complexes in a competitive setting: the small ST complexes are outcompeted by B before they can reach the critical size that would give them a fitness advantage.

Understanding the rates of complex formation, q and β, requires an understanding of how they depend on environmental conditions: complex formation can be favored in some environments and opposed in others. An example of ST where q depends on the environment is provided by the unicellular alga *Chlorella vulgaris*, which in the presence of a phagotrophic predatory flagellate forms colonies that are nearly invulnerable to predation (18). An example of CT for which β depends on the environment is provided by the bacterium *Bacillus subtilis*, which under certain stress conditions, like starvation or toxicity of the environment, forms biofilms. Biofilm formation is a developmental process in which bacteria undergo a regulated lifestyle switch from a nomadic unicellular state to a sedentary multicellular state, where subsequent growth results in structured communities and cellular differentiation (Kolter, 2010).

Furthermore, different environments can facilitate CT or ST. Bonner advanced the idea that ST is characteristic of multicellular forms of aquatic origin, whereas CT is typical in terrestrially derived lineages (Bonner, 1998). This could indeed be the case, because an aquatic environment with currents would make it difficult for units to find each other—thus β would be small and CT inefficient. In the absence of CT, however, ST is the natural solution. In terrestrial environments on the other hand, one can assume that the density of similar types on a patch of soil is high, which facilitates CT. In fact, in an environment where β is high and CT is easy to achieve, ST might have a hard time to compete. The reason is that complex formation via ST is slow when compared to CT in a situation where the local concentration of units is high.

5. Cooperation

Since both ST and CT require that several units function together as part of a complex, cooperation is proposed as a major factor in biological construction. We will show however that ST and CT represent two very different types of cooperation. Cooperation is always vulnerable to exploitation. The tension between cooperation and defection can be seen in our theory. Let us discuss this problem in the context of evolution of multicellularity. For ST a cell type could arise which does not contribute to the benefit of the organism but exploits it to enhance its own reproduction. Such a defecting, 'cancerous' cell undermines the advantages of multicellularity. For CT, however, the typical defector would be another cell type which participates in the aggregation of cooperators and subsequently exploits them (28). Here we find that our theory leads to basic aspects of evolutionary game dynamics; in this context, cooperation has been shown to evolve when certain mechanisms are in place (e.g. spatial structure, kin discrimination, group selection, reciprocity). For a discussion of these mechanisms see Nowak (2006). Thus, the emergence of multicellularity via CT requires a mechanism for the evolution of cooperation. Both ST and CT are discussed below.

5.1. ST

In the case of ST, there are many types of defection that we can imagine. Below we consider two simplified scenarios: an unproductive mutant which in no way affects the complex and a lethal mutant which kills the complex instantaneously. Any other type of mutation will be in between these two scenarios in terms of the amount of harm that it does to the complex.

5.1.1. Scenario 1: unproductive mutation

In this case, with probability u, a complex of size i produces a mutant which reverts to solitary life. It breaks from the complex and is unable to make complexes itself. In this case mutation is not harmful to the complex but it is unproductive in the sense that with a certain probability the complex produces solitary

individuals.

$$A_i \xrightarrow{a_i q(1-u)} A_{i+1}$$
$$A_i \xrightarrow{a_i(1-q)(1-u)} A_i + A$$
$$A_i \xrightarrow{a_i u} A_i + B$$
$$B \xrightarrow{b} B + B \tag{29}$$

We can describe these reactions by the following system of differential equations, where x_i is the abundance of complex A_i and y is the abundance of B.

$$\dot{x}_1 = (1-u)(1-q)\sum_i a_i x_i - q(1-u)a_1 x_1 - \phi x_1$$
$$\dot{x}_i = q(1-u)(a_{i-1}x_{i-1} - a_i x_i) - \phi x_i$$
$$\dot{y} = y(b-\phi) + u\sum_i a_i x_i \tag{30}$$

As before, ϕ is chosen such that $y + \sum i x_i = 1$. This implies that $\phi = by + \sum a_i x_i$. We find the equilibrium ϕ implicitly as

$$\sum_{i\geq 1}\prod_{j=1}^{i}\frac{q(1-u)a_j}{\phi + q(1-u)a_j} = \frac{q}{1-q} \tag{31}$$

The condition that x wins is $\phi > b$. For simplicity, let us assume $b=1$. Then $\phi > 1$ if

$$\sum_{i\geq 1}\prod_{j=1}^{i}\frac{q(1-u)a_j}{1 + q(1-u)a_j} > \frac{q}{1-q} \tag{32}$$

Let

$$g(u) = \sum_{i\geq 1}\prod_{j=1}^{i}\frac{q(1-u)a_j}{1 + q(1-u)a_j} - \frac{q}{1-q} \tag{33}$$

This is a decreasing function of u on the interval [0,1] and hence it will have at most one zero in the interval $0 \leq u \leq 1$. The function will have a zero in this interval if it has different signs when evaluated at 0 and 1 respectively. $g(1) = -q/(1-q) < 0$; now $g(0)$ gives the comparison between ST in the absence of mutation ($u=0$) and unicellularity. If $g(0) > 0$ then ST without mutation wins, otherwise unicellularity wins. Since $g(1) < 0$, the polynomial has one root in the interval [0,1] only if $g(0) > 0$. Thus we find that a necessary condition for ST with defection to be stable is that ST without defection is stable. Moreover, we find that the condition that ST wins over multicellularity is an error threshold condition: the mutation rate has to be lower than a certain threshold (see Fig. 6).

In particular, let us study the case where the maximum size of a complex is $n=2$. As before when studying what happens when a cutoff is imposed (i.e. $n=2$), the results obtained for the special case are not the same as for the general case, where there is no cutoff. In the case $n=2$ we find that ST is stable under defection only when

$$u < 1 - \frac{\sqrt{a_1^2(-1+2q)^2 + 4a_1a_2q} - a_1(-1+2q)}{2a_1a_2q} \tag{34}$$

Thus there exists an error threshold.

5.1.2. Scenario 2: lethal mutation

In this case, with probability u, a complex of size $i \geq 2$ produces a mutant which kills the entire complex. By itself, the mutant reproduces at rate b.

$$A_i \xrightarrow{a_i q(1-u)} A_{i+1}$$
$$A_i \xrightarrow{a_i(1-q)(1-u)} A_i + A$$
$$A_i \xrightarrow{a_i u} 0 \quad \forall i \geq 2$$

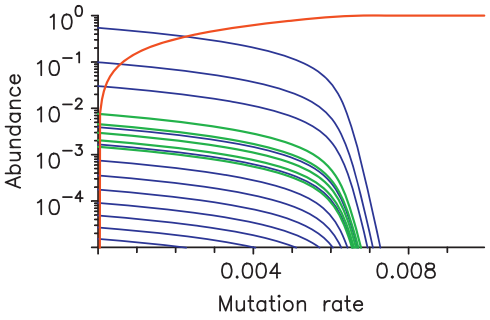

Fig. 6. An error threshold condition determines whether 'cooperators' or 'defectors' win in this ST model. We use the standard ST framework, but assume that an A unit can mutate into a B unit at rate u. The B unit leaves the complex and starts solitary reproduction. If the mutation rate is small enough, there is a mixed equilibrium (a mutation-selection balance). If the mutation rate exceeds a critical value, which here is approximately given by $u_c \approx 0.0081$, then A becomes extinct. Fitness landscape: $a_i = 0.95i$ for $i = 1,2,3$; $a_i = 2i$ for $i = 4,..,8$ and $a_i = 0.3i$ for all $i \geq 9$. The probability of staying together is $q=0.3$. Color code: green for x_i with $i=4..8$; blue for x_i with $i = 1,2,3,9,10\ldots$ (from top to bottom); red for y; equilibrium abundances are shown as function of the mutation rate. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

$$B \xrightarrow{b} B + B \tag{35}$$

We can describe these reactions by the following system of differential equations, where x_i is the abundance of complex A_i and y is the abundance of B.

$$\dot{x}_1 = (1-u)(1-q)\sum_i a_i x_i - q(1-u)a_1 x_1 - \phi x_1$$
$$\dot{x}_i = q(1-u)(a_{i-1}x_{i-1} - a_i x_i) - \phi x_i$$
$$\dot{y} = y(b-\phi) + u(1-q)\sum_i a_i x_i \tag{36}$$

As before, ϕ is chosen such that $y + \sum i x_i = 1$. This implies that $\phi = by + \sum a_i x_i$.

Note that the x_i equations in this case are the same as for the previous scenario. The only change is in the equations for y but these do not affect our calculation. Hence, in the case of a lethal mutation, we find the same condition (32) for multicellularity to be stable under invasion by a lethal mutant.

However, the case $n=2$ yields a different error threshold (because the system of differential equations is different when a cutoff is imposed), given by

$$u < 1 - \frac{1}{a_1(1+(-2+a_2)q)}. \tag{37}$$

5.2. CT defection

Let A and D be two types which can form aggregates either with their own type or with each other. The reactions below describe the formation of these aggregates. The relationships between the reaction rates will determine the relationships between A and D. In this section we want to consider the situation when D is a defector which takes advantage of A. We will describe below in what ways this can be achieved.

$$A_iD_j + A \xrightarrow{\beta} A_{i+1}D_j$$
$$A_iD_j + D \xrightarrow{\beta} A_iD_{j+1}$$
$$A_iD_j \xrightarrow{a_{ij}} A_iD_j + A$$
$$A_iD_j \xrightarrow{d_{ij}} A_iD_j + D$$
$$D \xrightarrow{b} 2D \tag{38}$$

D is a defector which exploits A as follows:

- the rate at which an A_iD_j complex produces more As is slower the more Ds there are in the complex ($a_{i,j} > a_{i,j+1}$ for all $i \geq 1$ and $j \geq 0$);
- the rate at which an A_iD_j complex produces more Ds is higher the more As there are in the complex ($d_{i,j} < d_{i+1,j}$ for all $i \geq 0$ and $j \geq 1$);
- the rate at which an A_iD_j complex produces As is slower than the rate at which it produces Ds ($d_{i,j} > a_{i,j}$ for all $i \geq 1, j \geq 1$).

One would expect such a defector D to outperform A. In such a situation, the survival of cooperation will depend on the existence of mechanisms that support the evolution of cooperation. Moreover, note that if we consider individual units as players in a game, then a complex of size i corresponds to a group of i individuals involved in a game interaction. Thus our theory naturally captures n-player games (Gokhale and Traulsen, 2010, 2011; Hauert et al., 2002, 2006; van Veelen and Nowak, 2012). It is usually the case in theory that interactions are instantaneous; here the complexes have certain lifetimes and hence individuals are involved in a particular interaction for a certain amount of time. This prevents them from being involved in other interactions.

If we moreover allow the existence of single units B which are unable to form complexes, then we are in the realm of n-player games with loners. Since B cannot participate to the formation of complexes, it means it cannot participate in an interaction and hence cannot play the game. For certain fitness landscapes, we can assume that A outperforms B (one requirement would be that $a_{i+1,0} \geq a_{i,0} > a_{1,0}$ for all $i \geq 2$) and B outperforms D. Thus A, D and B could be involved in a cyclic relationship which can allow the coexistence of the three types.

5.2.1. Example $n=2$

Let us consider a simple example of defection for the case in which complexes can reach at most size 2 and the defector D cannot form complexes with its own kind. Thus, D is an individual which can come together with an A individual and form a complex AD; this complex will only produce Ds. So D is a defector in the sense that it hijacks As ability to form complexes to produce more of itself. The reactions that describe the dynamics are

$$\begin{aligned}
&A \xrightarrow{a_1} A+A\\
&A+A \xrightarrow{\beta} AA\\
&AA \xrightarrow{a_2} AA+A\\
&A+D \xrightarrow{\beta} AD\\
&A \xrightarrow{a_3} AD+D\\
&D \xrightarrow{b} D+D
\end{aligned} \tag{39}$$

The equations that describe this dynamics are

$$\begin{aligned}
\dot{x}_1 &= a_1x_1 + a_2x_2 - 2\beta x_1^2 - \beta x_1 y - \phi x_1\\
\dot{x}_2 &= \beta x_1^2 - \phi x_2\\
\dot{x}_3 &= \beta x_1 y - \phi x_3\\
\dot{y} &= a_3x_3 + by - \beta x_1 y - \phi y
\end{aligned} \tag{40}$$

One question we want to ask is when y cannot invade the x equilibrium. For that, we calculate the x equilibrium as in Section 3.1 and ask when the linear system formed by x_3 and y has negative eigenvalues at that equilibrium. In other words, if we sprinkle some defectors at the equilibrium, we want them to not grow. We then find the condition for CT to be stable under invasion by the type of defector discussed above to be a lower bound on β. If β is greater than a certain function of the parameters of the model, then CT is stable.

6. Discovery of a new niche

So far we have only analyzed the case where the solitary units and the complexes compete for the same ecological niche. However, one of the advantages of complex formation is the possibility to explore new niches. For example, in the case of multicellularity, an increase in size would allow an escape from phagotrophic predation (Boraas et al., 1998; Stanley, 1973); but even in the absence of predation, increased size could provide advantages such as the discovery of novel metabolic opportunities (Pfeiffer and Bonhoeffer, 2003). Moreover, multicellularity can allow the emergence of cellular differentiation and division of labor. Some cells in the multicellular complex could adopt different functions and thereby make it possible for the organism to explore new ecological niches. This can have a profound impact. For example, we argued above that if there is only one niche, then the choice of $a_i = ib$ for the reproductive rates is neutral. However, one can easily envisage a situation where complexes of size i are as productive as i solitary individuals, but they manage to inhabit a new ecological niche and are therefore selected. For example, the freshwater green alga *Scenedesmus acutus* is unicellular in lab cultures, but if exposed to water that contains its predator *Daphnia*, it starts to form colonies by staying together—when it divides, it retains the daughter cells within the cell walls. The multicellular alga grows and photosynthesizes at the same rate as the unicells, but it sinks more rapidly, thereby avoiding predation (Grosberg and Strathmann, 2007).

In this section we investigate the possibility that large complexes discover a new niche. This means that from a certain size upwards, complexes discover a new niche which changes the nature of the competition with B, which is inhabiting the old niche. We will analyze this scenario for both ST and for ST with chain breaking and conclude that it has the interesting property that rather inefficient complex organisms can successfully compete with their solitary ancestors, if those are limited to the original niche. The complex organisms would lose the competition in the original niche, but prevail over all by discovering the new niche. In this setting coexistence between complex and solitary is possible (Fig. 7).

6.1. ST

The equations that describe the discovery of a new niche in the case of ST are

$$\begin{aligned}
\dot{x}_1 &= (1-q)\sum_{i=1}^{\infty} a_i x_i \psi_i - q a_1 x_1 \psi_1 - d x_1\\
\dot{x}_i &= q(a_{i-1}x_{i-1}\psi_{i-1} - a_i x_i \psi_i) - d x_i \quad i = 2,3,\ldots\\
\dot{y} &= y(b\psi_0 - d)
\end{aligned} \tag{41}$$

where

$$\psi_i = \begin{cases} 1 \Big/ \left[1 + \eta \sum_{j=1}^{n} j x_j\right] & \text{for } i = 0,1,\ldots,n\\ 1 \Big/ \left[1 + \eta \sum_{j=n+1}^{\infty} j x_j\right] & \text{for } i = n+1, n+2,\ldots \end{cases} \tag{42}$$

is the density limitation. The ψ determines who competes with whom: the first n complexes compete with B because they are in the same niche, whereas the complexes from size $n+1$ upwards are in a niche of their own.

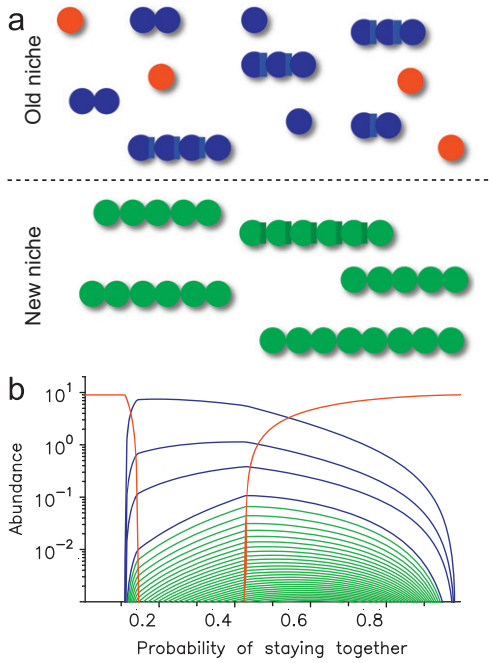

Fig. 7. Discovery of a new niche via staying together. (a) Complexes of size $i=1,\ldots,4$ compete with the unicellular ancestor in the original niche. Complexes of size $i\geq 5$ populate a new niche (with separate density regulation). (b) Simulation results: the evolutionary outcome depends on the probability of staying together, q. The curves from top to bottom represent the equilibrium abundances of x_1 through x_n. The blue curves are the complexes that share the same niche as the solitary ancestors (red); the green curves are the complexes that have discovered a new niche. We obtain the following equilibria: for $q<0.11$ all-B is stable, for $0.11<q<0.16$ there is a mixed equilibrium, for $0.16<q<0.42$ all-A is stable, for $0.42<q<0.98$ there is again a mixed equilibrium between A and B, for $0.98<q$ all-B is stable. In this ST model coexistence is possible because of the presence of two niches. Fitness landscape: $a_i=0.8i$ for $i=1,..,50$ and $a_i=40$ for all $i\geq 50$. Thus A has an intrinsically lower reproduction rate than B; it is only selected because of the new niche. Color code: blue for x_i with $i=1..4$ and green for x_i with $i\geq 5$ (both are ordered from top to bottom); red for y. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Notice that in this section we omit the kinetic equations. This is because they are identical to the ones for ST. The only way in which the discovery of a new niche is reflected is in the density limitation and that is not represented in the kinetic equations.

We performed simulations (Fig. 7) for the fitness landscape: $b=1$, $a_i=0.8i$ for $i\leq 50$ and $a_i=40$ for $i>50$. Thus, for this landscape, the first 50 complexes are worse than singletons and all others from size 51 onwards have constant reproductive rate. We find that there exist 5 regions depending on the rate q of complex formation. If q is too small, then complexes never grow fast enough to discover the new niche. If q is too large then very large complexes are being produced, which is inefficient (because after size 50, increasing the size does not yield more productivity). Intermediate values of q allow for either the dominance of ST or for a mixed equilibrium between ST and singletons. The coexistence is possible because of the two niches.

6.2. A simple analytical model

We can analytically study a simplified system where complexes only reach size 2. Here A_1 and B compete for the old niche while A_2 discovers a new niche.

$$\dot{x}_1 = a_1x_1(1-2q)\psi_1 + a_2x_2\psi_2 - dx_1$$

$$\dot{x}_2 = a_1qx_1\psi_1 - dx_2$$

$$\dot{y} = y(b\psi_1 - d) \tag{43}$$

where $\psi_1=1/(1+\eta(x_1+y))$ and $\psi_2=1/(1+\eta x_2)$.

Without loss of generality, let us assume $b=1$. We want to find the regions of mixed equilibrium; solving for the equilibrium, we impose the condition that $x_1>0$, $x_2>0$ and $y>0$ and find these to be equivalent to $q_0<q<q_1$ and $q_2<q<1$ where

$$q_0 = \frac{d(1-a_1)}{a_1(a_2-2d)} \tag{44}$$

and q_1 and q_2 are the two roots of the quadratic equation

$$2a_1^2(1-d)q^2 + qa_1(1-a_1+d+a_1d-a_2) + d(1-a_1) = 0 \tag{45}$$

All three critical values exist and are positive if $a_1<1$ and $a_2>2$. In this case we find only three critical values (hence only four parameter regions). These regions are the same as in our general simulations above, except that we are missing the last region—for this simple model large values of q still allow coexistence because, unlike before, large inefficient complexes do not exist (complexes of size 2 are not inefficient).

7. Conclusion

In summary, we have proposed a simple mathematical framework that allows us to study the evolution of construction in biology. Our theory rests on two basic operations, staying together and coming together, which represent different routes for the evolution of complexity. We argue that these operations are crucially involved at all levels of construction – in the formation of protocells, eukarya, multi-cellular organisms and animal societies – and moreover that they are very distinct mechanisms facing distinct challenges. These differences become very clear when we discuss the problem of cooperation and defection. In both operations we feel the tension between 'the constructive capabilities of cooperation and the dissipative forces of selfishness' (Krakauer, 2011) but they manifest themselves in different form. ST leads to linear selection equations and the condition for evolution of cooperation is a mutation threshold. CT is an evolutionary game and a specific mechanism for the evolution of cooperation is needed. In both cases the complexes that are being formed can exploit new niches and evolve division of labor.

Both ST and CT have been discussed previously, but they have been treated within the same multilevel selection approach (Michod, 2007). We suggest that there is a significant difference between these two operations and what can be achieved through them. True CT means that the units that are coming together could be genetically diverse (not of a clonal origin), which immediately leads to a conflict of interest (cooperation versus defection). This conflict can be resolved through multilevel selection or other mechanisms for the evolution of cooperation (Nowak, 2006). In contrast, ST leads to genetic diversity only because of mutations that might occur during the growth of the aggregate. In the absence of these mutations the daughter cells are not independent units with independent interests that need to be aligned, and consequently there is no multilevel selection for ST. Thus, the origin of ST, in the sense of competing A (which does ST) versus B (which is solitary) is not a multi-level selection problem. Similarly, competing different ST strategies is again not a multi-level selection problem.

Finally, we propose that CT without ST does not lead to the formation of higher-level individuals. ST is the essential mechanism for achieving this step. An organism that always needs CT in order to form multicellularity can never fully resolve the tension arising from different types, with different interests. What is ultimately needed for the evolution of complex multicellularity is

for CT to be replaced by ST. An interesting complication arises if one considers that sexual reproduction can be seen as a form of CT. Sexual reproduction has the typical CT features of generating novel outcomes (creativity) and leading to tensions - as in game dynamical situations (Haig, 2002; Hofbauer and Sigmund, 1998). On the other hand sexual reproduction is a special type of CT, because two cells fuse and give rise to one unit, which helps to suppress conflicts of interests. An empirical classification of various types of CT and ST, how they are employed for construction and what can be achieved with either of them is an interesting avenue for future research.

Acknowledgments

We are indebted to Julien Ayroles, Nicole King, Sarah Kocher, Roberto Kolter, Richard M. Losick, Richard E. Michod, Robert M. Pringle and David Smyth for valuable discussions and comments which have significantly improved our manuscript. Support from the John Templeton Foundation and the NSF/NIH joint program in mathematical biology (NIH Grant R01GM078986) is gratefully acknowledged.

Appendix A. Exploring different density limitations and death rates

Our choice ϕ for the density limitation affects the death rate of complexes. Another option is to consider a density limitation that affects the birth rates, similar to the one we introduced for the discovery of a new niche. Thus consider the following system:

$$\dot{x}_1 = (1-q)\sum_i a_i\psi x_i - qa_1\psi x_1 - d_1 x_1$$

$$\dot{x}_i = q(a_{i-1}\psi x_{i-1} - a_i\psi x_i) - d_i x_i \quad i = 2,3,\ldots$$

$$\dot{y} = y(b\psi - d) \tag{46}$$

where $\psi = 1/[1+\eta(y+\sum_i i x_i)]$ and η is a constant. The density limitation in this system becomes a limitation on the birth rate rather than on the death rate; the birth rate for each complex is multiplied by the same constant ψ. The larger the population, the fewer individuals of each type will be born. We find ψ implicitly as before to be given by

$$\frac{1-q}{q}\sum_{i\geq 1}\prod_{j=1}^{i}\frac{\psi q a_j}{d_j+\psi q a_j} = 1 \tag{47}$$

The condition that A wins is equivalent to $\psi < d/b$ which becomes

$$\frac{1-q}{q}\sum_{i\geq 1}\prod_{j=1}^{i}\frac{\frac{d}{b}qa_j}{d_j+\frac{d}{b}qa_j} > 1 \tag{48}$$

In particular, if $d_i = d$ for all i, then the condition that A wins is

$$\frac{1-q}{q}\sum_{i\geq 1}\prod_{j=1}^{i}\frac{qa_j}{b+qa_j} > 1 \tag{49}$$

which is the same as condition (6) found for the ϕ density limitation. Thus, using the ψ density limitation leads to the same competitive exclusion conclusion and for $d_i = d$, it leads to the same condition. Of course, other density limitations can and should be explored, especially when the framework is applied to study a particular system; but we hope that this analysis shows that at least some of the main conclusions are unaffected by the choice of density limitation.

Appendix B. Neutrality for $a_i = ib$

In our discussion of ST we argued that the fitness landscape $a_i = ib$ is neutral; there we defined neutrality to mean that $\phi = b$ which we showed to hold for any y and x_i if $a_i = ib$. Here we make that statement stronger, by showing that we can actually find a family of equilibria for which $\phi = 1$. We will do the calculation for ST with chain breaking because it is more interesting, but the same principle applies for all types of ST.

Write $x_k = u_k x_1$ for all $k > 1$ and $x_1 = x$. Moreover, write $\overline{u} = \sum_{k\geq 2} u_k$. Then the equilibrium conditions for the system (12) are

$$1 = c\overline{u}$$

$$3(1+c)u_2 = 1 + 2c\overline{u}$$

$$(1+c)(k+1)u_k = (k-1)u_{k-1} + 2c\left(\overline{u} - \sum_{2\leq j<k} u_j\right) \tag{50}$$

The solution to this system can be easily shown by induction to be $u_k = (1/(1+c))^{k-1}$ for $k \geq 2$. To see that this is so, note that the middle equation from above gives $u_2 = 1/(1+c)$. Assuming that $u_k = (1/(1+c))^{k-1}$ for all $k < n$ with $n \geq 3$, then from the third equation above we find

$$\begin{aligned} u_n &= \frac{n-1}{n+1}\left(\frac{1}{1+c}\right)^{n-1} + \frac{2}{n+1}\frac{1}{1+c}\left(1 - c\sum_{1\leq k<n-2}\left(\frac{1}{1+c}\right)^{k}\right) \\ &= \frac{n-1}{n+1}\left(\frac{1}{1+c}\right)^{n-1} + \frac{2}{n+1}\left(\frac{1}{1+c}\right)^{n-1} = \left(\frac{1}{1+c}\right)^{n-1} \end{aligned} \tag{51}$$

It is easy to check that $\phi = \sum_k a_k x_k = \sum k(1/(1+c))^{k-1} = 1$. Thus, we have found a family of equilibria $x_k = u_k x$ which are neutral.

References

Alexander, R.D., 1974. The evolution of social behavior. Annu. Rev. Ecol. Syst. 5, 325–383.
Bell, G., 1997. Size and complexity among multicellular organisms. Biol. J. Linn. Soc. 60, 345–363.
Bonner, J.T., 1988. The Evolution of Complexity by Means of Natural Selection. Princeton University Press, Princeton.
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.
Boraas, M.E., Seale, D.B., Boxhorn, J.E., 1998. Phagotrophy by a fagellate selects for colonial prey: a possible origin of multicellularity. Evol. Ecol. 12, 153–164.
Carroll, S.B., 2001. Chance and necessity: the evolution of morphological complexity and diversity. Nature 409, 1102–1109.
Chen, I.A., Roberts, R.W., Szostak, J.W., 2004. The emergence of competition between model protocells. Science 305, 1474–1476.
Crespi, B.J., 2008. Social conflict resolution, life history theory, and the reconstruction of skew. In: Jones, C., Hager, R. (Eds.), Reproductive Skew in Vertebrates. Cambridge University Press.
Crow, J.F., Kimura, M., 2009. An Introduction to Population Genetics. Blackburn Press.
Doebeli, Ispolatov, 2010. Complexity and diversity. Science 328, 494–497.
Eigen, M., Schuster, P., 1977. The hyper cycle. A principle of natural self-organization. Part A: emergence of the hyper cycle. Naturwissenschaften 64, 541–565.
Ewens, W.J., 2010. Mathematical Population Genetics. Springer.
Fontana, W., Buss, L.W., 1994. The arrival of the fittest: toward a theory of biological organization. Math. Biol. B 56, 1–64.
Furusawa, C., Kaneko, K., 2000. Complex organization in multicellularity as a necessity in evolution. Artif. Life 6, 265–281.
Gadagkar, R., 2001. The Social Biology of Ropalidia Marginata: Toward Understanding the Evolution of Eusociality. Harvard University Press, Cambridge.
Gadagkar, R., Bonner, J.T., 1994. Social insects and social amoebae. J. Biosci. 19, 219–245.
Gokhale, C., Traulsen, A., 2010. Evolutionary games in the multiverse. Proc. Natl. Acad. Sci. USA 107, 5500–5504.
Gokhale, C., Traulsen, A., 2011. Strategy abundance in evolutionary many-player games with multiple strategies. J. Theor. Biol. 238, 180–191.
Grosberg, R.K., Strathmann, R.R., 2007. The evolution of multicellularity: a minor major transition. Annu. Rev. Ecol. Evol. Syst. 38, 621–654.
Gueron, S., 1998. The steady-state distributions of coagulation–fragmentation processes. J. Math. Biol. 37, 1–27.

Gueron, S., Levin, S.A., 1995. The dynamics of group formation. Math. Biosci. 128, 243–264.
Gueron, S., Levin, S.A., Rubenstein, D.I., 1996. The dynamics of herds: from individuals to aggregations. J. Theor. Biol. 182, 85–98.
Haig, D., 2002. Genomic Imprinting and Kinship. Rutgers University Press, New Brunswick, New Jersey.
Hall-Stoodley, L., Costerton, J.W., Stoodley, P., 2004. Bacterial biofilms: from the natural environment to infectious diseases. Nature 2, 95–108.
Hauert, Ch., De Monte, S., Hofbauer, J., Sigmund, K., 2002. Volunteering as red queen mechanism for cooperation in public goods game. Science 296, 1129–1132.
Hauert, C., Michor, F., Nowak, M.A., Doebeli, M., 2006. Synergy and discounting of cooperation in social dilemmas. J. Theor. Biol. 239 (2), 195–202.
Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge.
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.
Hunt, J.H., 2011. A conceptual model for the origin of worker behaviour and adaptation of eusociality. J. Evol. Biol. , http://dx.doi.org/10.1111/j.1420-9101.2011.02421.x.
King, N., 2004. The unicellular ancestry review 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. Comp. Biol. 43, 247–253.
Kirk, D.L., 2005. A twelve-step program for evolving multicellularity and a division of labor. BioEssays 27, 299–310.
Knoll, A.H., 2011. The multiple origins of complex multicellularity. Annu. Rev. Earth. Planet Sci. 39, 217–239.
Kolter, R., 2010. Biofilms in lab and nature: a molecular geneticist's voyage to microbial ecology. Int. Microbiol. 13, 1–7.
Krakauer, D., 2011. Laws of cooperation. Science 332, 538–539.
Krebs, J.R., Davies, N.B. (Eds.), 1991. Behavioural Ecology: An Evolutionary Approach, third edn. Blackwell Scientific Publications.
Leadbeater, E., Carruthers, J.M., Green, J.P., Rosser, N.S., Field, J., 2011. Nest inheritance is the missing source of direct fitness in a primitively eusocial insect. Science 333, 874–876.
Lynch, M., Conery, J.S., 2003. The origins of genome complexity. Science 302, 1401–1404.
Margulis, L., 1981. Symbiosis in Cell Evolution. Life and its Environment on the Early Earth. Freeman, San Francisco.
May, R.M., 2001. Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton.
Maynard Smith, J., Szathmary, E., 1998. The Major Transition in Evolution. Oxford University Press, Oxford, UK.
Michod, R.E., 2007. Evolution of individuality during the transition from unicellular to multicellular life. Proc. Natl. Acad. Sci. 104, 8613–8618.
Michod, R.E., Roze, D., 1999. Cooperation and conflict in the evolution of individuality. III. Transitions in the unit of fitness. In: Nehaniv, C.L. (Ed.), Mathematical and Computational Biology: Computational Morphogenesis, Hierarchical Complexity, and Digital Evolution. Lectures on Mathematics in the Life Sciences, vol. 26; 1999, pp. 47–91.
Michod, R.E., Roze, D., 2001. Cooperation and conflict in the evolution of multicellularity. Heredity 86, 1–7.
Michener, C.D., 1974. The Social Behavior of the Bees. Harvard University Press, Cambridge, MA.
Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.
Nowak, M.A., Tarnita, C.E., Wilson, E.O., 2010. The evolution of eusociality. Nature 466, 1057–1062.
Okubo, A., 1986. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. Adv. Biophys. 22, 1–94.
Pfeiffer, T., Bonhoeffer, S., 2003. An evolutionary scenario for the transition to undifferentiated multicellularity. Proc. Natl. Acad. Sci. 100, 1095–1098.
Queller, D.C., 1997. Cooperators since life began. Q. Rev. Biol. 72, 184–188.
Queller, D.C., 2000. Relatedness and the fraternal major transitions. Philos. Trans. Biol. Sci. 355, 1647–1655.
Rainey, P.B., Kerr, B., 2010. Cheats as first propagules: a new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. BioEssays 32, 872–880.
Rossetti, V., Schirrmeister, 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, 23–24.
Rossetti, V., Filippini, M., Svercel, M., Barbour, A.D., Bagheri, H.C. Emergent multicellular life cycles in filamentous bacteria owing to density-dependent population dynamics. J.R.Soc. Interface, 10.1098/rsif.2011.0102, in press.
Stanley, S.M., 1973. An ecological theory for the sudden origin of multicellular life in the Late Precambrian. Proc. Natl. Acad. Sci. USA 70, 1486–1489.
van Veelen, M., Nowak, M.A., 2012. Multi-player games on the cycle. J. Theor. Biol. 292, 116–128.
Webb, J.S., Givskov, M., Kjelleberg, S., 2003. Bacterial biofilms: prokaryotic adventures in multicellularity. Current 6, 578–585.
Willensdorfer, M., 2008. Organism size promotes the evolution of specialized cells in multicellular digital organisms. J. Evol. Biol. 21, 104–110.
Wolpert, L., 1990. The evolution of development. Biol. J. Linn. Soc. 39, 109–124.