*International Journal of Astrobiology*

**cambridge.org/ija**

# A thermodynamic limit constrains complexity and primitive social function

Philip J Gerrish[1,2,3] and Claudia P Ferreira[4]

[1]School of Biology, Georgia Tech, 310 Ferst Dr NW, Atlanta, GA 30332, USA; [2]Theoretical Biology & Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA; [3]Instituto de Ciencias Biomédicas, Universidad Autónoma de Ciuadad Juarez, 32310 Chihuahua, México and [4]Departamento de Bioestatística, Instituto de Biociências, Universidade Estadual Paulista, 18618-000 Botucatu, São Paulo, Brazil

**Research Article**

**Cite this article:** Gerrish PJ, Ferreira CP (2019). A thermodynamic limit constrains complexity and primitive social function. *International Journal of Astrobiology* **18**, 329–335. https://doi.org/10.1017/S1473550418000149

Received: 12 December 2017
Revised: 28 March 2018
Accepted: 3 April 2018
First published online: 13 June 2018

**Key words:**
Cancer; complexity; cooperation; extraterrestrial life; game theory; multicellularity; phase transition; social evolution

**Author for correspondence:**
Philip J. Gerrish, E-mail: pgerrish@gatech.edu

## Abstract

The evolutionary trend toward increasing complexity and social function is ultimately the result of natural selection's paradoxical tendency to foster cooperation through competition. Cooperating populations ranging from complex societies to somatic tissue are constantly under attack, however, by non-cooperating mutants or transformants, called 'cheaters'. Structure in these populations promotes the formation of cooperating clusters whose competitive superiority can alone be sufficient to thwart outgrowths of cheaters and thereby maintain cooperation. But we find that when cheaters appear too frequently – exceeding a threshold mutation or transformation rate – their scattered outgrowths infiltrate and break up cooperating clusters, resulting in a cascading loss of social cohesiveness, a switch to net positive selection for cheaters and ultimately in the loss of cooperation. Our findings imply that a critically low mutation rate had to be achieved (perhaps through the advent of proofreading and repair mechanisms) before complex cooperative functions, such as those required for multicellularity and social behaviour, could have evolved and persisted. When mutation rate in our model is also allowed to evolve, the threshold is crossed spontaneously after thousands of generations, at which point cheaters rapidly invade. Probing extrapolations of these findings suggest: (1) in somatic tissue, it is neither social retro-evolution alone nor mutation rate evolution alone but the interplay between these two that ultimately leads to oncogenic transitions; the rate of this coevolution might thereby provide an indicator of lifespan of species, terrestrial or not; (2) the likelihood that extraterrestrial life can be expected to be multicellular and social should be affected by ultraviolet and other mutagenic factors.

## Introduction

Speculations about life on other planets (Levin *et al.*, 2017), taking their lead from observations of our own planet, almost invariably suppose that extra-terrestrial life will have some degree of complexity (cf. (Ward and Brownlee, 2000; Lane and Martin, 2010)). Complexity is ultimately the product of cooperation: two or more 'units' coming together because of the benefits derived from their cooperation. Indeed, it is hard to imagine life without the complexity and, while not technically a requirement for 'life' under its common definitions, the tight association between complexity and life would suggest that the former is almost an immediate derivative of the latter. At least on our planet, it would seem that cooperative interaction among individual parts or among individuals must have followed soon after the advent of self-replication (Maynard Smith and Szathmáry, 1995; Queller, 1997).

Such interaction, however, often carries energetic costs and there is thus an intrinsic evolutionary advantage associated with *not* interacting. The astonishing predominance of socially cohesive groups and multicellular organisms on our planet has therefore presented evolutionary biologists with an intriguing theoretical puzzle: to explain the evolutionary emergence of positive social interaction, or cooperation (Nowak, 2012). Evolutionary implications of this puzzle are far from trivial, as it has direct bearing on most of the 'major transitions' in evolution (Maynard Smith and Szathmáry, 1995; Queller, 1997; Szathmáry, 2015; West *et al.*, 2015; Levin *et al.*, 2017). Solutions to this puzzle invoke different mechanisms that include group selection (Nowak and Sigmund, 1998), direct (Axelrod and Hamilton, 1981; Nowak and May, 1992) and indirect (Nowak and Sigmund, 1998) reciprocity, tagging (Riolo *et al.*, 2001) and kin selection (Hamilton, 1964). Group selection is likely to be effective only under special circumstances (Price, 1972). Reciprocity, tagging and sophisticated kin selection require recognition systems that are often complex and are thought to be themselves products of a more primitive form of kin selection (Axelrod and Hamilton, 1981): individuals simply help their neighbours who, due to the population's spatial structure, are likely to be kin. And by helping kin, they increase the survival probability of their lineage. In this framework, it is easy to see how cooperative traits might be favoured, offering probably the most

© Cambridge University Press 2018

CAMBRIDGE UNIVERSITY PRESS

parsimonious explanation for the primordial origins of social behaviour in its broadest sense and at least partial explanations for the subsequent evolution of more complex innovations such as social groups (Hamilton, 1964), multicellularity (Michod and Roze, 2001), growth restraint (Merlo *et al.*, 2006), rudimentary moral systems (Alexander, 1987) and virulence modulation (Frank, 1994).

Not only is cooperation ubiquitous on our planet, but the physiological apparatuses associated with cooperation are numerous and can be quite complex. Indeed, a significant fraction of an organism's genome may be devoted to such apparatuses (Santorelli *et al.*, 2008). Consequently, a significant fraction of all mutations (or transformations) may hinder cooperation – a notion corroborated by experimental data (Velicer *et al.*, 1998, 2000; Strassmann *et al.*, 2000; Santorelli *et al.*, 2008). Most such cooperative functions carry an energetic cost and mutations that hinder these functions would, therefore, allow more energy to be devoted to individual growth and replication. It seems likely that such mutations, therefore, would often carry an immediate selective advantage at the individual level – another notion corroborated by experimental data (Velicer *et al.*, 1998, 2000; Turner and Chao, 1999; Strassmann *et al.*, 2000). Furthermore, these mutants would continue to enjoy the benefits derived from neighbours whose cooperative functions remain intact. In sum, by neglecting their dues to society, these mutants can further their individual success and are therefore called *cheaters* (Velicer, 2003; Velicer and Vos, 2009). The ready availability of cheater mutations and the threat they pose to populations has been demonstrated quite strikingly in laboratory microbial populations (Turner and Chao, 1999; Fiegna and Velicer, 2003; Harrison and Buckling, 2005) and by models (Nelson and Perelson, 1995; Nowak and Sigmund, 2004) like our mean-field first model (Methods). These observations lead to a new puzzle: given the very reasonable and recently substantiated premises that social cheater mutations are common and carry an immediate selective advantage, how can cooperation persist? We investigate the most rudimentary mechanism to promote the persistence of cooperation, namely population structure (Nowak *et al.*, 1994) and we find there exists a critical rate of production of cheater mutants above which this mechanism fails.

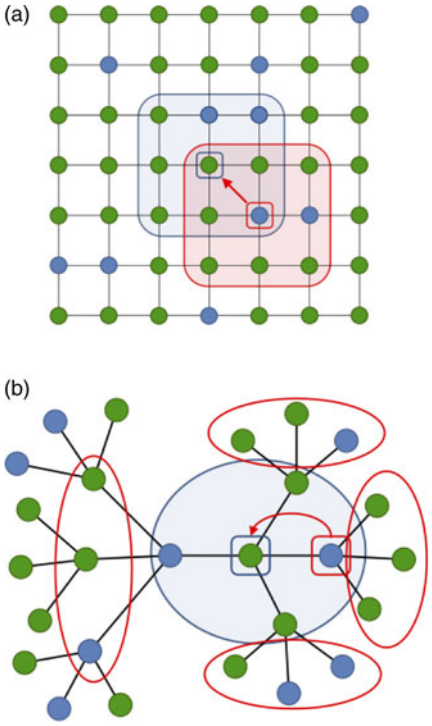



**Fig. 1.** Two different network configurations. Green nodes represent 'social' individuals; blue nodes represent 'asocial' individuals. Focal nodes are indicated by a small blue square; a candidate displacement comes from an immediate neighbour indicated by a small red square. Immediate neighbours of the focal individual are surrounded by a blue line; the focal individual's fitness increases as the number of immediate neighbours increase. Immediate neighbours of the candidate displacing neighbour are surrounded by a red line; the candidate displacing neighbour has a fitness that is also an increasing function of the number of *its* neighbours who are cooperators. To analytically model the evolution of cooperation through local competition on such networks, we employed the *pair-approximation* technique (Baalen, 2000; Ellner, 2001) (see SI). (a) a regular lattice network; (b) a regular tree network.

The most closely related precedents to our investigation (Ohtsuki *et al.*, 2006; Taylor *et al.*, 2007; Ichinose *et al.*, 2018) have focused mainly on the evolutionary emergence of cooperation rather than its maintenance, but some maintenance-related results have been obtained there. Within its broader context, this previous work addressed the question of whether an anomalous cheater mutant (or 'defector') introduced into a population of cooperators is likely to reach fixation and thus displace the cooperating wildtype. This approach is the standard population genetics approach but, while related, it does not naturally extrapolate to questions of cooperation maintenance. In particular, it does not take into account the synergistic effect of other cheater mutants appearing elsewhere in the population. The biological evidence (Velicer *et al.*, 1998; Strassmann *et al.*, 2000, 2008; Velicer *et al.*, 2000) would suggest that cheater mutations are not anomalous, isolated events but in fact can be quite common and are thus more accurately modelled as recurring at some rate throughout the population. Other related precedents have looked at how mutation may erode cooperation in iterative games when it is asymmetric (Garcia and Traulsen, 2012) or 'local' (of small effect on strategy) versus non-local (Stewart and Plotkin, 2015).

## Modeling the evolution of cooperation with mutation

To model the maintenance of cooperation, we start with a population of cooperators and allow cheaters to arise by mutation (Gerrish, 2010; Allen *et al.*, 2012). We suppose that the cooperating wildtype helps other individuals in the population by increasing their fitness by an amount $\varepsilon$. If cooperation is altruistic, individuals only help others but do not help themselves; if cooperation is mutualistic, as in the production of a locally-shared public good, individuals help themselves as much as they help others (West *et al.*, 2006). Also, we suppose that cheater mutants may receive help from cooperating wildtype but do not themselves help others; the energy and/or replication time they save by not helping others increases their individual fitness by an amount $\beta$. Cooperators become cheaters by mutation at rate $\mu$ (forward rate) and cheaters become cooperators at rate $\mu'$ (backward rate). Heritable cooperation is typically the result of genetically programmed cooperative functions, while heritable cheating is typically due to the loss of such functions. For genetic reasons, therefore, it is reasonable to assume that $\mu > \mu'$ .

In an unstructured population, dynamics are well-approximated by mean-field equations in which an individual may interact with any other individual in the population with equal probability (mass action). Under these well-mixed conditions, any amount of

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press

mutation results in the loss of altruistic cooperation for any $\varepsilon > 0$ and $\beta > 0$ or in the loss of mutualistic cooperation for $\beta > \varepsilon$.

In a structured population, an individual's interactions are limited to a 'circle of friends', which may be modelled as neighbours of that individual on the population's interaction network which we assume to be static (c.f., Pacheco *et al.*, 2006). Such a network is reflective of the kind of population structure, such as surface adhesion, that most likely gave rise to the earliest manifestations of social behaviour. In our model of structured populations, each individual occupies a node on a network and has $n$ immediate neighbours on average. Figure 1 gives a synopsis of our model; details are found in the online supplementary material.

## Results

Our simulations show that, when interactions within a population are structured along the links of a network, primitive kin selection can effectively preserve cooperation despite recurrent cheater-mutations. The clustering necessary for kin selection to be effective, however, can be maintained and cooperation thus preserved, only as long as the mutation rate is low enough. There exists a threshold mutation rate (Fig. 2) above which cooperating clusters are disrupted from within by the advantageously scattered (Fig. 3) appearance and expansion of cheater mutants: as neighbourhoods increasingly consist of cheaters, support from neighbours decreases, kin selection erodes and it no longer pays to cooperate. As a result, neighbourhoods increasingly consist of cheaters and the cycle continues. This positive feedback results in a cascading loss of social cohesiveness above the threshold mutation rate – resembling a second-order phase transition – and ultimately in the complete loss of cooperation.

Using analytical methods in the supplementary material, we derive an expression for the threshold cheater mutation rate above which cooperation cannot persist: $\mu_c \approx k(\alpha + 1)(\phi\varepsilon - \beta)$, where $\phi = 1$ if cooperation is altruistic and $\phi = 2$ if cooperation is mutualistic, $\alpha$ is an index of how deterministic selection is (typically $\alpha = 1$) and $k = 0.030$, a constant of proportionality (see

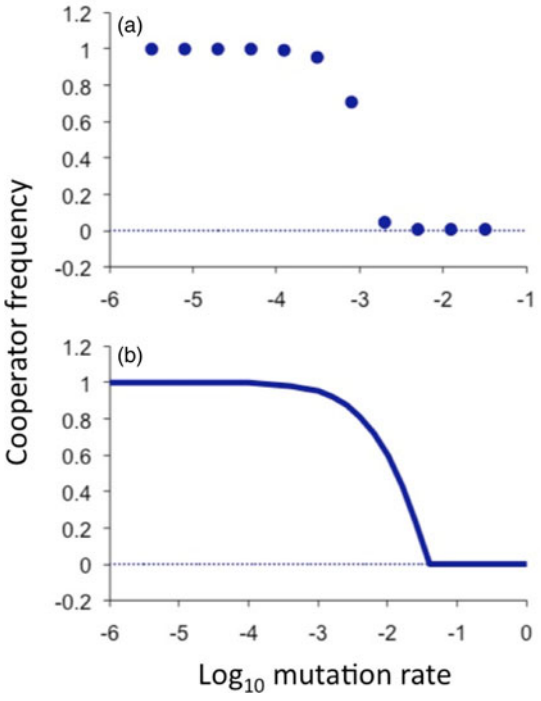



**Fig. 2.** Equilibrium cooperator frequencies as a function of mutation rate. (a) As determined from simulated populations that were spatially structured on a three-dimensional grid. The population initially consists of all cooperators and cheaters arise by spontaneous mutation. The intrinsic advantage of being a cheater (or cost of being a cooperator) was $\beta = 0.01$, the advantage conferred by a cooperating neighbour was $\varepsilon = 0.05$ and total population size was 27 000. (b) As predicted by the analytical equilibrium solution of the 'pair approximation' model (solving the equations $f_p = 0$ and $f_q = 0$ for $p$ and $q$, as defined in the supplementary material and reporting $p$), using the same parameter values but implicitly assuming an infinite population. The threshold calculated by this model is higher than the threshold observed in simulations because of well-known error in this approximation scheme due to stochastic demographic effects. Nevertheless, the correction factor is constant over a wide range of parameter values and across various network topologies (supplementary material). This analytical curve reveals a sharp transition between the maintenance of cooperation and the complete loss of cooperation; the sharpness of the inflection resembles that of a second-order phase transition (supplementary material).

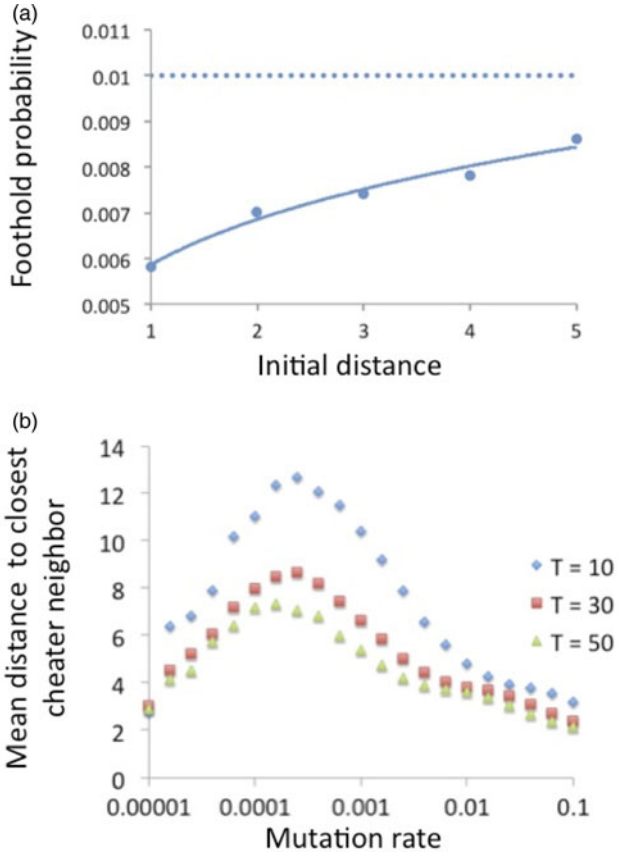



**Fig. 3.** Early growth dynamics. (a) Probabilities that cheaters gain a foothold in a structured population as a function of the initial distance between cheaters; 'foothold probability' is defined as the probability of a cheater being fixed in a small fraction of the population in the absence of mutation. Here, three cheaters were placed at random on a $10 \times 10$ grid such that each cheater was initially separated from the other two cheaters by a given distance (horizontal axis). Fixation probabilities (vertical axis) were calculated as the fraction of times cheaters reached fixation in a total of 50 000 simulations. This is not an individual-based fixation probability as typically calculated in population genetics but a collective fixation probability. The dotted line shows the probability of fixation of a neutral mutation; the fact that foothold probabilities are less than this value reflects the fact that in the absence of recurrent mutation, cheaters are not expected to be successful. (b) The average minimum distance between a cheater and its nearest cheating neighbour as a function of mutation rate. These averages were taken over all cheaters during the first 10, 30 and 50 generations of growth. At small mutation rates, rare cheater presence is due mostly to local replication, hence the small distance between cheaters and their nearest neighbours. As mutation rate increases, this distance increases as scattered mutants contribute more and more to the mean; eventually, scattered cheaters permeate the population enough to make this distance begin to decline.

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press

supplementary material). From this expression, we derive the condition that $\mu_c > 0$ only if $\varepsilon > \beta$ when cooperation is altruistic, meaning that some level of cheater mutation can be tolerated only if the benefit conferred by a cooperating neighbour exceeds the benefit of being a cheater; this is a restatement of Hamilton's rule (Hamilton, 1964) for clonal organisms in which relatedness approaches unity (West *et al.*, 2006).

Cheaters have a fitness advantage when they are very rare (Fig. 4(a)) because most of them are isolated and surrounded by cooperators (Fig. 3(a)). As their frequency increases, however, their average fitness declines because as they grow, they leave offspring at neighbouring nodes (i.e., they form cheating clusters); on average, they become surrounded by more fellow cheaters and less cooperators and their average fitness decreases accordingly. On a global scale, therefore, cheaters become less fit than their cooperating counterparts as they grow (Fig. 4(a,c)). This means that if cheaters are to reach fixation – as they do when $\mu > \mu_c$ – then they must cross an adaptive valley (Fig. 4(a)) which is problematic because it violates natural selection's tendency to increase fitness. The problem is resolved, however, by looking at the fitness dynamics of individuals relative only to their immediate neighbours (Fig. 4(b)). On this local scale, at high mutation rates ($\mu > \mu_c$), cheaters dominate their immediate surroundings at all frequencies, starting with a significant selective advantage when rare and decreasing in fitness as they grow until reaching a relative fitness of one at fixation. At low mutation rates ($\mu < \mu_c$), cheater relative fitness is lower overall on a global scale (Fig. 4(c)) because their small presence is due less to mutation and more to rare outgrowths (Fig. 3(b)), resulting in fewer clusters of larger average size (Fig. 4(e)); on average, cheaters are surrounded by more cheaters – a fact that is also reflected by the lower perimeter-to-area ratio of cheating clusters when mutation rates are low ($\mu < \mu_c$) (Fig. 4(f)). Figure 4(d) reveals the decisive consequence of this qualitative difference in cluster morphology: when $\mu < \mu_c$, local relative fitness of cheaters dips below one as cheater frequency increases – a hallmark of the game of Chicken (Velicer, 2003), which is known to maintain cooperation in a population because of the devastating costs of mutual (or collective) defection, as happens within the larger cheater clusters. The foregoing observations shed light on spontaneous growth patterns in structured populations, how these patterns selectively thwart cheater outgrowths when $\mu < \mu_c$ and how these patterns change qualitatively as mutation rate exceeds $\mu_c$, abruptly removing the selective repression of cheater growth.

The observed interplay between mutation rate and cooperation leads one naturally to question what happens when both are evolving, as is surely the case in real populations. When we allow both to evolve in our simulations, we find that there is a long period of apparent stasis followed by a relatively sudden rise of cheaters to fixation (Fig. 5(a)). This intriguing dynamic appears to be the inevitable result of genetic linkage. Mutators (individuals with elevated mutation rates) are more likely to produce cheater mutants than non-mutators and there is thus an intuitive association between cheaters and mutators. If mutators remain linked to the cheaters they produce then, initially, there is indirect selection for mutators because the cheaters they produce have an advantage when rare. Mutators increasingly permeate the population and cheater presence thus becomes more scattered, as evidenced by the increasing number of clusters and decreasing cluster size (Fig. 5(b)). Eventually, cheaters reach a critical rate of production throughout the population (through increased mutator presence) and cooperating clusters suddenly become ineffective against the spread of cheaters. At this point, cheaters invade (Fig. 5(a,b)) and concurrently: (1) average mutation rate increases (Fig. 5(c)), reflecting the fact that mutators 'hitchhike' with cheaters to high frequency (Sniegowski *et al.*, 1997) and (2) fitness decreases (Fig. 5(d)), reflecting the triumph of defectors in the game of Prisoner's Dilemma (Nowak and Sigmund, 2004) and the related Tragedy of the Commons (Hardin, 1968).

## Discussion

The ubiquity of cooperation in nature and the cohesive complexity it produces indicates cooperation's robustness as an evolutionary strategy. The collapse of cooperation can be costly to a population in both relative and absolute senses: it can put a population at a disadvantage relative to other populations in which cooperation remains intact and it can impose an absolute cost by decreasing the intrinsic viability of a population even to the point of extinction. Identifying mechanisms that can lead to the collapse of cooperation, therefore, is of interest to both basic and applied science. Here, we have identified one such mechanism, namely, mutation rate and its evolution. Others have explored how asymmetry in mutation rates can erode cooperation in iterative games (Garcia and Traulsen, 2012) and have found that unconstrained evolution of the magnitudes of costs and benefits of cooperation in addition to strategies can lead to the collapse of cooperation (Stewart and Plotkin, 2014, 2015).

Our threshold result raises the intriguing hypothesis that primitive life forms, that likely resided on surfaces giving them a population structure ideal for the emergence of cooperation, nevertheless had to achieve critical replication fidelity before social behaviour and multicellularity, as evolutionary innovations, could have appeared. A simple organism with a high mutation rate could not have maintained social functions, despite any population structure it may have, whereas a more complex organism with a lower genomic mutation rate, on the other hand, would have been capable of acquiring new social genes. In related simulations, we have found that newly acquired social functions cannot invade a structured population when the mutation rate is high, but do eventually invade when mutation rate is low. These findings would further suggest that, were extraterrestrial life to be found, the likelihood it would be multicellular and social should be affected by the replication fidelity it was able to achieve. High replication fidelity could be difficult to achieve in severe environments, such as might be found on other planets where radiation and other mutagenic factors might abound, and our findings would thus predict limited complexity and social evolution of life forms that might exist there.

Our findings might, in a very parsimonious way, shed some light on early evolutionary processes leading up to the transition from normal to cancerous tissue. Normal somatic cells are genetically programmed to make enormous evolutionary sacrifices at the level of the individual cell, for the common good of the somatic population (Merlo *et al.*, 2006); knockouts of any aspect of this programming can result in one or more of the six hallmarks of cancer cells (Hanahan and Weinberg, 2000), most of which are characteristic of cheaters. Other evidence shows a relation between some cancers and elevated somatic mutation rates (Modrich, 1994). Our results would suggest: (1) that somatic tissue with simple population structure can tolerate a surprisingly high rate of appearance of potentially cancerous cheater mutants thereby providing the first line of defense prior to effective immune surveillance and (2) that it may be the interplay between

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press

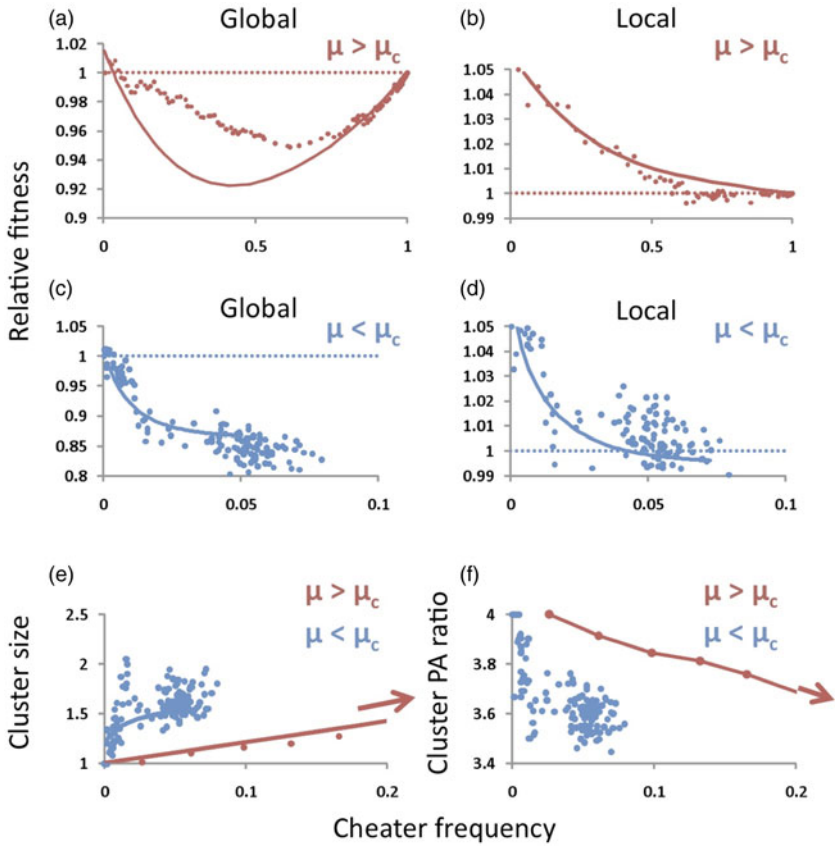



**Fig. 4.** Understanding the threshold. Solid lines plot the 'pair approximation' analytical predictions and dots plot simulation results. Parameters were: $\varepsilon = 0.1$, $\beta = 0.01$, $\alpha = 2$, $n = 8$ and $\mu = 0.001$ for the low mutation rate ($\mu < \mu_c$, blue lines) and $\mu = 0.01$ for the high mutation rate ($\mu > \mu_c$, red lines). For these parameter values, our theory predicts $\mu_c \approx 0.008$. (a,b) Global (a) and local (b) relative fitnesses for a mutation rate exceeding the critical value. (c,d) Global (c) and local (d) relative fitnesses for a mutation rate below the critical value. (e,f) Cluster size as cheater frequency increases (e). When mutation rate is low (blue line), cheaters exist in larger clusters, on average, with smoother peripheries evidenced by lower perimeter-to-area (PA) ratio (f). Within these clusters, fitness is low because of lack of mutual support; this situation mimics the over-consumption of a common resource and the 'tragedy of the commons' that results (West *et al.*, 2006), apparent in panel (d) where local relative fitness dips below one.

somatic cheating and mutation rate evolution that eventually leads to the transition from normal somatic tissue to cancerous tissue.

## Methods

### *Mean field model*

Mean-field dynamics of cooperating wildtype and cheater frequencies are given by the replicator equations $dx/dt = x(r_x - r_x X - r_y Y) + \mu' y - \mu x$ and $dy/dt = y(r_y - r_x X - r_y Y) + \mu x - \mu' y$, respectively, where $r_x = \varepsilon x$ and $r_y = \beta + \varepsilon x$ for altruistic cooperation and $r_x = \varepsilon + \varepsilon x$ and $r_y = \beta + \varepsilon x$ for mutualistic cooperation. In this analysis, we consider both the limit $\mu' \to 0$ as well as the case $\mu = \mu'$. When cooperation is altruistic under these unstructured conditions, cheaters always displace the wildtype for all $\varepsilon > 0$ and $\beta > 0$ $\hat{x} = 0$ and $\hat{y} = 1$ for the case $\mu' \to 0$; $\hat{x} \approx (\mu/\beta)$ and $\hat{y} \approx 1 - (\mu/\varepsilon)$ for the case $\mu = \mu'$. When cooperation is mutualistic, cheaters displace the wildtype if $\beta$

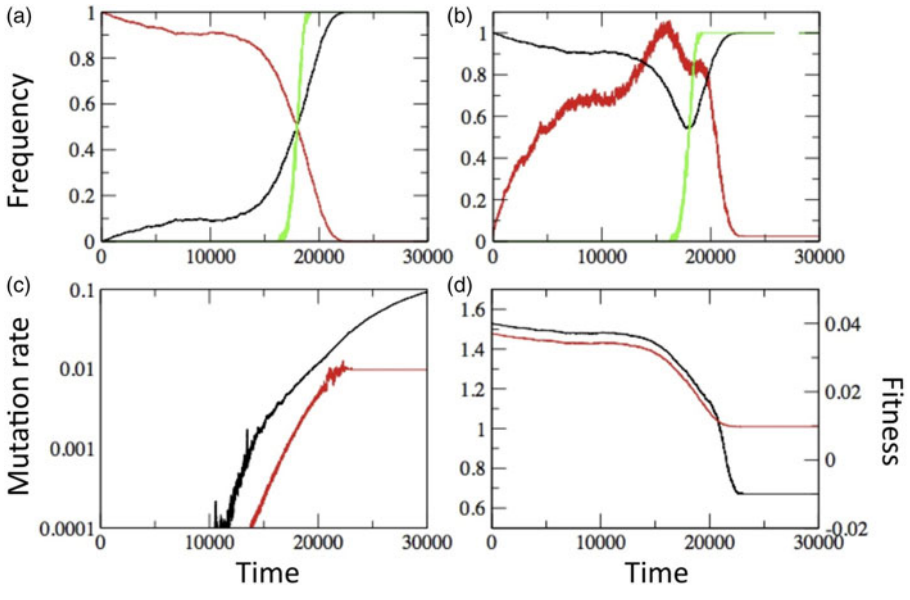



**Fig. 5.** Results of simulations in which both cheaters and mutation rate were allowed to evolve. The initial mutation rate was $10^{-5}$ cheater mutations per replication. Parameters values were: $\varepsilon = 0.05$, $\beta = 0.01$, $\alpha = 2$, $n = 8$. (a) Frequency of cooperators (red) and cheaters (black) as a function of time. The green curve shows the probability that the major cluster is formed by cheaters, alluding to the related concepts of a percolation threshold and phase transition phenomena. (b) Number of clusters (red) and major cluster size (black); the two measures are normalized, respectively, by 40 (maximum) and 2500 (lattice size). Again, the green curve shows the probability that the major cluster is formed by cheaters. (c) Cooperator (red) and cheater (black) average mutation rate. (*D*) Mean fitness of cooperators and cheaters (red and black). All curves show the means of 100 simulations and lattice size of $50 \times 50$.

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press

$> \varepsilon$: $\hat{x} = 0$ and $\hat{y} = 1$ for $\mu' \to$ ; $\hat{x} \approx (\mu/(\beta - \varepsilon))$ and $\hat{y} \approx 1 - (\mu/(\beta - \varepsilon))$ for $\mu = \mu'$. The wildtype displaces cheaters if $\varepsilon > \beta$: $\hat{x} = 1$ and $\hat{y} = 0$ for $\mu' \to 0$ and $\hat{x} \approx 1 - (\mu/(\varepsilon - \beta))$, $\hat{y} \approx (\mu/(\varepsilon - \beta))$ for $\mu = \mu'$.

### *Structured-population model*

Simulations proceed in discrete time with synchronous updating. The fitness of an individual is increased by the presence of each cooperating neighbour on the network by a factor of $1 + \varepsilon$. If the individual in question is a cheater, its fitness is further increased by a factor of $1 + \beta$. An individual's fitness is therefore calculated as $(1 + \beta)^a \ (1 + \varepsilon)^s$, where $a = 1$ if the individual is a cheater ('asocial') and $a = 0$ if the individual is a cooperator and $s$ denotes the number of cooperating ('social') immediate neighbours it has. The highest possible fitness is thus awarded to a cheater whose immediate neighbours are all cooperators. Let $w_f$ denote the fitness of a 'focal' individual and let $w_i$ denote the fitness of one of the immediate neighbours of the focal individual. The probability that this neighbour's offspring will displace the focal individual is defined as: (1) the probability that the focal individual will leave no offspring thus leaving its node vacant, $e^{-w_f}$, times (2) the probability that the neighbour in question will outcompete the other neighbours for that vacancy, $w_i^\alpha / \sum_{j \in \Omega f} w_j^\alpha$, where $\Omega_f$ denotes the neighbourhood of the focal individual and $\alpha$ denotes the sampling exponent: we use $\alpha = 2$ in figures which is appropriate for bacterial or somatic populations – it reflects the probability that both daughter cells survive sampling; sampling becomes more deterministic as $\alpha$ increases. The analytical theory (supplementary material) mimics the simulations and employs the 'pair-approximation' technique. The analytical theory is applicable to a wide range of networks and random, scale-free and regular networks were all implemented in simulations.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S1473550418000149

**Acknowledgements.** We especially thank G. Velicer for guidance and careful reading. We thank C. Maley, A. Colato, P. Sniegowski, G. Fernandez, R. Lenski, I. Gordo, L. Perfeito, M. VanBaalen, V. Souza and L. Eguiarte for discussions and helpful comments. This work was mainly supported by grants from NIH (R01 GM079483 and ARRA Supplement to P.J.G) and National Aeronautics and Space Administration (NNA15BB04A to P.J.G.); support was also received from the European Commission (FP7 231807 to P.J.G.) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 16/23738-3 to C.P.F.). Part of this work was done during a CNRS-funded visit (P.J.G.) to the Laboratoire Jean Kuntzmann, University of Grenoble Alpes, France.

**Author contributions.** PG conceived and directed this study, performed mathematical analyses and wrote the paper; CF and PG designed and ran simulation studies; CF aided in the writing/editing of the paper and SI. Both authors gave final approval for publication.

We have no competing interests.

## References

**Alexander RD** (1987) *The Biology of Moral Systems*. New York: Aldine de Gruyter.

**Allen BA, Traulsen A, Tarnita CE and Nowak MA** (2012) How mutation affects evolutionary games on graphs. *Journal of Theoretical Biology* **299**, 97–105.

**Axelrod R and Hamilton WD** (1981) The evolution of cooperation. *Science* **211**, 1390–1396.

**Baalen V** (2000) Pair approximations for different spatial geometries. In Dieckmann U, Law R and Metz JAJ (eds), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, pp. 359–387.

**Ellner SP** (2001) Pair approximation for lattice models with multiple interaction scales. *Journal of Theoretical Biology* **210**(4), 435–447.

**Fiegna F and Velicer GJ** (2003) Competitive fates of bacterial social parasites: persistence and self-induced extinction of *Myxococcus xanthus* cheaters. *Proceedings of the Royal Society of London B* **270**, 1527–1534.

**Frank SA** (1994) Kin selection and virulence in the evolution of protocells and parasites. *Proceedings. Biological Sciences/The Royal Society* **258**, 153–161.

**Garcia J and Traulsen A** (2012) The structure of mutations and the evolution of cooperation. *PloS ONE* **7**(4), e35287.

**Gerrish PJ** (2010) Mutation rate and the maintenance of cooperation: a parsimonious model of somatic evolution and oncogenic transitions. *Proceedings of Computational and Mathematical Methods in Science and Engineering* **2010**, 1129–1132.

**Hamilton WD** (1964) The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**, 1–16.

**Hanahan D and Weinberg RA** (2000) The hallmarks of cancer. *Cell* **100**(1), 57–70.

**Hardin G** (1968) The tragedy of the commons. *Science* **162**, 1243–1248.

**Harrison F and Buckling A** (2005) Hypermutability impedes cooperation in pathogenic bacteria. *Current Biology* **15**, 1968–1971.

**Ichinose G, Satotani Y and Sayama H** (2018) How mutation alters the evolutionary dynamics of cooperation on networks. *New Journal of Physics*. http://iopscience.iop.org/10.1088/1367-2630/aac2a7

**Lane N and Martin W** (2010) The energetics of genome complexity. *Nature* **467**, 929.

**Levin SR, Scott TW, Cooper HS and West SA** (2017) Darwin's aliens. *International Journal of Astrobiology* **69**, 1–9.

**Maynard Smith J and Szathmáry Er** (1995) *The Major Transitions in Evolution*. Oxford: Oxford University Press.

**Merlo LM, Pepper JW, Reid BJ and Maley CC** (2006) Cancer as an evolutionary and ecological process. *Nature Reviews Cancer* **6**(12), 924–935.

**Michod RE and Roze D** (2001) Cooperation and conflict in the evolution of multicellularity. *Nature Heredity* **86**, 1–7.

**Modrich P** (1994) Mismatch repair, genetic stability and cancer. *Science* **266** (5193), 1959–1960.

**Nelson GW and Perelson AS** (1995) Modeling defective interfering virus therapy for AIDS: conditions for DIV survival. *Mathematical Biosciences* **125** (2), 127–153.

**Nowak MA** (2012) Why we help: the evolution of cooperation. *Scientific American* **307**(1), 34–39.

**Nowak MA and May RM** (1992) Evolutionary games and spatial chaos. *Nature* **359**, 826–829.

**Nowak MA and Sigmund K** (1998) Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577.

**Nowak MA and Sigmund K** (2004) Evolutionary dynamics of biological games. *Science* **303**(5659), 793–799.

**Nowak MA, Bonhoeffer S and May RM** (1994) Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences of the USA* **91**, 4877–4881.

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

**Pacheco JM, Traulsen A and Nowak MA** (2006) Coevolution of strategy and structure in complex networks with dynamical linking. *Physical Review Letters* **97**(25), 258103.

**Price GR** (1972) Extension of covariance selection mathematics. *Annals of Human Genetics* **35**, 485–490.

**Queller DC** (1997) Cooperators since life began. *The Quarterly Review of Biology* **72**(2), 184–188.

**Riolo RL, Cohen MD and Axelrod R** (2001) Evolution of cooperation without reciprocity. *Nature* **414**, 441–443.

**Santorelli LA, Thompson CRL, Villegas E, Svetz J, Dinh C, Parikh A, Sucgang R, Kuspa A, Strassmann JE, Queller DC and Shaulsky G**

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press

(2008) Facultative cheater mutants reveal the genetic complexity of cooperation in social amoebae. *Nature* **451**(7182), 1107–1110.

**Sniegowski PD, Gerrish PJ and Lenski RE** (1997) Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**(6634), 703–705.

**Stewart AJ and Plotkin JB** (2014) Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences of the USA* **111**, 17558–17563.

**Stewart AJ and Plotkin JB** (2015) The evolvability of cooperation under local and non-local mutations. *Games* **6**, 213–250.

**Strassmann JE, Zhu Y and Queller DC** (2000) Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. *Nature* **408**, 965–967.

**Szathmáry E** (2015) Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences* **112**, 10104–10111.

**Taylor PD, Day T and Wild G** (2007) Evolution of cooperation in a finite homogeneous graph. *Nature* **447**(7143), 469–472.

**Turner PE and Chao L** (1999) Prisoner's dilemma in an RNA virus. *Nature* **398**, 441–443.

**Velicer GJ** (2003) Social strife in the microbial world. *TRENDS in Microbiology* **11**, 330–337.

**Velicer G and Vos M** (2009) The evolution of cooperation in the Myxobacteria. *Annual Review of Microbiology* **63**, 599–623.

**Velicer GJ, Kroos L and Lenski RE** (1998) Loss of social behaviors by *Myxococcus xanthus* during evolution in an unstructured habitat. *Proceedings of the National Academy of Sciences of the USA* **95**, 12376–12380.

**Velicer GJ, Kroos L and Lenski RE** (2000) Developmental cheating in the social bacterium *Myxococcus xanthus*. *Nature* **404**, 598–601.

**Ward PD and Brownlee D** (2000) *Rare Earth: Why Complex Life is Uncommon in the Universe*. New York: Copernicus Books.

**West SA, Fisher RM, Gardner A and Kiers ET** (2015) Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences of the USA* **112**(33), 10112–10119.

**West SA, Griffin AS, Gardner A and Diggle SP** (2006) Social evolution theory for microorganisms. *Nature Review Microbiology* **4**(8), 597–607.

https://doi.org/10.1017/S1473550418000149 Published online by Cambridge University Press