Competing metabolic strategies in a multilevel selection model

André Amado¹, Lenin Fernández¹, Weini Huang², Fernando F. Ferreira³ and Paulo R. A. Campos¹

¹Evolutionary Dynamics Lab, Department of Physics, Federal University of Pernambuco, 50670-901 Recife, Pernambuco, Brazil
²Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany
³Center for Interdisciplinary Research on Complex Systems, University of São Paulo, 03828-000 São Paulo, Brazil

The evolutionary mechanisms of energy efficiency have been addressed. One important question is to understand how the optimized usage of energy can be selected in an evolutionary process, especially when the immediate advantage of gathering efficient individuals in an energetic context is not clear. We propose a model of two competing metabolic strategies differing in their resource usage, an efficient strain which converts resource into energy at high efficiency but displays a low rate of resource consumption, and an inefficient strain which consumes resource at a high rate but at low yield. We explore the dynamics in both well-mixed and structured populations. The selection for optimized energy usage is measured by the likelihood that an efficient strain can invade a population of inefficient strains. It is found that the parameter space at which the efficient strain can thrive in structured populations is always broader than observed in well-mixed populations.

1. Introduction

The handling of a common resource by individuals with different means to exploit the resource often gives rise to social dilemmas [1]. Here social dilemma is referred to as the situation in which individuals benefit from selfishness unless everyone chooses the same alternative. One of the well-known examples is observed when these features are related to how rapidly and how efficiently the resource is exploited. With the massive available data from experiments, especially in microbial populations, the social dilemma ensuing from resource competition has been more effectively addressed [2]. The conflict arises directly from the
trade-off between growth rate and yield of resource exploitation. Empirical studies demonstrate that the trade-off between resource uptake and yield is commonplace in the microbial world [3–6]. The trade-off owes to biophysical limitations which prevent organisms from optimizing multiple traits simultaneously. Independent experiments have reported the existence of a negative correlation between rate and yield [3,7–9]. The existence of the trade-off for adenosine triphosphate (ATP)-producing reactions can be derived using arguments from irreversible thermodynamics [2,10].

Important insights have been gained regarding the understanding of energy efficiency at the cell standpoint. By energetic efficiency we mean the ability of cells to extract energy for a given amount of resource. It was shown that gene expression across the whole genome typically changes with growth rate [11]. This evidence naturally emerges from the assumption of occurrence of different trade-offs at the intracellular level [11,12], regardless of the metabolic pathway used to convert resource into ATP. The pathway choice dictates the speed and efficiency of growth. Heterotrophic organisms are good examples of the yield–rate trade-off. Their metabolism converts resource like glucose into energy in the form of ATP, mainly through two processes: fermentation and respiration [13,14]. Fermentation does not require oxygen and uses glucose as a reactant to produce ATP. However, it is considered to be inefficient because the end products still contain a great deal of chemical energy. On the other hand, in respiration the glucose undergoes complete oxidation. It is efficient because several ATPs are produced from a single glucose molecule [15]. A microorganism using fermentation depletes the resource faster than microorganisms using respiration [13–15]. However, respiration allows cells to make a better use of resource and to produce more offspring given the same amount of resource. In a population of cells using different metabolic processes, there will be competition between highly efficient cells (respiration) and inefficient cells (fermentation).

This competition among traits presenting distinct metabolic pathways is the concern within a broad evolutionary perspective. The physiological states of a cell play an important role. At the physiological level, previous investigations have tried to understand the underlying conditions that may trigger the switch from one mode of metabolism to another [6]. The most influential factor studied so far is the resource influx rate into the system. This problem has been addressed in the framework of spatially homogeneous environments [16], as in spatially structured environments of unicellular organisms [5,17]. The studies of the steady state of population dynamics were carried out either by writing down a set of coupled differential equations describing population density and resource dynamics or by directly mapping the problem into a Prisoner’s Dilemma game, thus requiring some arbitrary definition of a pay-off matrix in terms of the resource influx rate. Conclusions can then be drawn from the established results in the field of evolutionary game theory. According to those studies, competition always drives the strain with the lowest rate of ATP production to extinction in homogeneous environments [5]. On the other hand, when space is structured, the outcome of the system is determined by its free energy content (resource influx rate). If the resource is abundant, the strain using the pathway with high rate and lower yield dominates. If the resource is limited, the strain using the pathway with high yield but at lower growth rate dominates [5].

Despite the attempts to study this trade-off on structured populations, these works have focused on spatially structured populations [5,17]. While this provides a valuable approach to study biological structures like biofilms, it lacks a clear definition of group and group reproduction, essential features of multicellular organisms as they are usually defined in animals, plants and fungi [18]. In this sense, these models describe more a multi-organism community rather than a multicellular organism. It is known that multicellular life can appear without a biofilm-like phase, for example, through the experimental evolution of multicellularity [19]. Our approach does not depend explicitly on spatial structure, instead we focus on the group structure. This way, we are able to study multicellular structures that can reproduce as groups. It is also critical to understand which evolutionary force can turn these multicellular structures into evolutionary stable units.

This work aims to investigate the conditions under which the optimized handling of resource can prevail when selection acts at different levels of biological organization. As such, a model of structured populations with multiple groups is presented. We assume explicit competition among individuals for resource. Although resource competition occurs at the individual level, the groups are assumed to grow and split at rates which depend on composition, putting this model within the framework of multilevel selection. Therefore, a higher level of organization is formed and the group itself could be seen as a replicating entity, a premise for the establishment of a multicellular organism. The issue especially motivated as multicellularity is conjectured to have arisen almost concomitantly with the great oxygenation event, and thereby allowing the evolution of the respiration mode of cellular metabolism.
Here, the problem is mainly addressed within a stochastic context whereby an evolutionary invasion analysis is considered. We do extensive computer simulations, though analytical developments are also carried out in the special case of well-mixed populations. A great limitation of using deterministic approaches to study population dynamics [5] is that the neglected stochastic effects may be critical in the evolution of natural populations, especially in small populations, a very likely situation if the resource is not abundant. Although this problem has been mapped into a Prisoner’s Dilemma game with constant pay-offs [16,17], it is based on a simplified assumption of linear interactions among individuals. Instead, we avoid these limitations by making a thorough exploration of the cellular properties and measuring their effects on the population dynamics. This is certainly a gap in this literature, as pointed out by MacLean [2].

The paper is organized as follows. In §2, we present the details of our model and simulations. In §3, we derive the equilibrium solutions, perform an invasion analysis and explain the simulation results. In §4, we give a detailed discussion of our results.

2. Methods

2.1. Resource-based model description

We consider a structured population with groups. The population consists of individuals which use either of the two competing metabolic strategies: (i) efficient use of resource, denoted by $C$, at which an individual converts resource into ATP efficiently and thereby achieves a high yield and (ii) rapid metabolism, hereby denoted by $D$, with which an individual consumes resource at a high rate but converts it into ATP inefficiently.

The number of groups $N_G$ and their local population sizes $P_i$ ($i = 1, \ldots, N_G$) vary over time. The influx of resource to the system (e.g. glucose) as well as how the resource is used are critical for the population dynamics. The population size increases or shrinks according to the amount of resource provided [20,21]. In every time step, individuals can produce new individuals and groups can split into new groups [22,23]. Each individual divides into two identical cells as soon as its internal energy storage $E_{ij}$ (group index, $i = 1, \ldots, N_G$, and individual index, $j = 1, \ldots, P_i$) surpasses the energy threshold $E_{\text{max}}$. Each daughter cell inherits half of the energy of the parental cell. Groups have a threshold size $P_{\text{max}}$, above which a group will split into two with its members being evenly distributed between the two new groups. In every time step, each individual dies with a constant probability $\nu$, i.e. every time step and for each individual a random number $x$ is drawn from a uniform distribution, $x \in [0, 1]$, and if $x$ is smaller than $\nu$, the individual dies. Both group splitting and the death process introduce stochasticity into the system. A group disappears when the number of individuals in that group shrinks to zero. The above-described dynamical rules are also illustrated in figure 1.

The amount of resource available to the population is constant and equal to $S$ in every time step. We consider equipartition of resource among groups and thus the amount of resource per group is $S_G = S/N_G$. Note that the total number of groups $N_G$ is a dynamical number. We break the complete process of an individual gaining energy into two stages: first, individuals compete for the resource; second, the caught resource is converted into internal energy (ATP). The assumption of these two stages is in accordance with mechanistic models of cells [11]: first, an enzyme transports the resource into the cell; then, a second set of enzymes metabolizes the resource into a generic form of energy, which includes ATP. The rates of resource import and of metabolism are distinct [11]. The metabolism of a cell is subject to multiple levels of regulation [24–26]. It is known that nutrients are used as substrates for growth but the nutritional state also supplies signals for the cell [27]. For example, cAMP (cyclic adenosine monophosphate) plays a central role in the regulation of the response of Escherichia coli to different nutritional states [28], and the design of the cell is an evolutionary choice that tunes it to act effectively within its biological niche. Note that our model does not refer to the resolution of processes at the intracellular level.

2.1.1. Implementing the resource uptake

The resource is shared among the members of a group. As a strain type with rapid metabolism, $D$, has a higher rate of consumption, it seizes a larger portion of the resource available to the group compared with strains type $C$, characterized by a low consumption rate. The amount of resource per strain $D$ in a
element split
when its internal energy surpasses $E_{\text{max}}$ the cell
divides into two new daughter cells
dead
proliferation

Death

Group split

Group split occurs when group
population surpasses $P_{\text{max}}$

number of cells

Figure 1. A graphical representation of the processes of group splitting, cell division and death.

given group $i$ is estimated as

$$S_D^i(S_G) = \frac{P_D^i(S_G)}{P_D^i(S_G) + P_C^i(S_G)} S_G,$$

(2.1)

where $P_D^i$ and $P_C^i$ are, respectively, the number of individuals for type $D$ and $C$ in group $i$, such that

$P_D^i + P_C^i = P^i$ corresponds to the size of group $i$. The consumption rates for strain type $C$ and $D$, $f_C^i(S_G)$ and $f_D^i(S_G)$, are functions of $S_G$, the amount of resource per group. Similarly, the amount of resource per strain $C$ is

$$S_C^i(S_G) = \frac{f_C^i(S_G)}{f_D^i(S_G) P_D^i + f_C^i(S_G) P_C^i} S_G.$$

(2.2)

Therefore, the total amount of resource for group $i$ is

$S_D^i(S_G) P_D^i + S_C^i(S_G) P_C^i = S_G$ as declared above.

2.1.2. Implementing the conversion of resource into internal energy

The resource seized by the cell is used to increment its internal energy $E_{ij}$, which is calculated as

$$\Delta E_{ij} = J_{\text{ATP}}^C(S_C^i) \Delta t \quad \text{or} \quad \Delta E_{ij} = J_{\text{ATP}}^D(S_D^i) \Delta t,$$

(2.3)

depending on the strain type, $C$ or $D$. The functions $J_{\text{ATP}}^C(S_C^i)$ and $J_{\text{ATP}}^D(S_D^i)$ represent how efficiently

the resource ($S_C^i$ or $S_D^i$) is transformed into energy (ATP) for each individual type. In a discrete-time model as assumed here, $\Delta t = 1$. The units of the defined variables/parameters are given in the electronic supplementary material.

2.2. Simulation procedure

2.2.1. Statistical measurement

In a pool of organisms using resource inefficiently, it is very likely that an alternative metabolic pathway

arises by mutation. It is important to evaluate the probability that a single individual of type $C$ invades

and finally takes over the whole population. The simulations start from an isogenic population of

strains $D$, which later evolves to a stationary regime. At this steady state, the population size is

determined by the population dynamics and the resource availability. For this reason, the system is said

to be self-organized. Once the homogeneous population evolves to its stationary states, a single strain $C$

is introduced to replace a randomly chosen cell from the population. From this time on, the fate of the

mutant type $C$ is observed, i.e. the simulation is tracked until the mutant gets lost or fixed.

The fixation probability is simply the fraction of independent runs at which the efficient strain gets

fixed. Because the population size itself depends on the parameters of the model, it is more reasonable

to use the relative fixation probability defined as the absolute fixation probability divided by $1/N_{\text{st}}$. Here

$N_{\text{st}}$ stands for the population size of strains $D$ at the moment the efficient strain is introduced. Thus,

$1/N_{\text{st}}$ corresponds to the fixation probability of a mutant under neutral selection. If the relative fixation

probability is larger than 1, the mutant type is said to be selected for, while a probability smaller than 1

means the strain is counter-selected.
2.2.2. Parametrization of the model

The rates of resource uptake and the rate of converting resource into energy appearing in equations (2.1) and (2.2) are defined as

\[
\dot{J}_C^S(S_C) = A_C, \quad (2.4a)
\]

\[
\dot{J}_D^S(S_C) = A_D, \quad (2.4b)
\]

\[
\dot{J}_C^{\text{ATP}}(S_C^i) = A_C^{\text{ATP}} \left(1 - \exp(-a_C^{\text{ATP}}S_C^i) \right), \quad (2.4c)
\]

and

\[
\dot{J}_D^{\text{ATP}}(S_D^i) = A_D^{\text{ATP}} \left(1 - \exp(-a_D^{\text{ATP}}S_D^i) \right). \quad (2.4d)
\]

The \(J^{\text{ATP}}\) functions display a sigmoidal shape, thus saturating for large values of \(S\), similar to Michaelis–Menten functions [5,11]. The effects of the rates on the population dynamics are extensively studied here. For most of the parameters, we perform a sweep over the parameter space that covers the physically meaningful regions. The maximum values for the rates are chosen in such way that the rate–yield trade-off exists. By definition, the efficient strain \(C\) is the one that transforms resources into internal energy efficiently at the expense of the consumption rate, while strain \(D\) metabolizes fast but at low yield. The parameters above follow some conditions:

(i) \(A_D\) must be larger than \(A_C\) \((A_D > A_C)\), which ensures that type \(D\) always has a higher consumption rate for the same amount of resource. As naturally emerging in the forthcoming calculations, the relevant parameter is in fact the ratio between these two quantities, \(\epsilon = A_D/A_C\).

(ii) Another condition is \(\Delta_{\text{ATP}} = a_D^{\text{ATP}}/a_C^{\text{ATP}} < 1\). Again, the ratio between the exponents in the functions (2.4c,d) is a key quantity in the process. This condition warrants that strain \(D\) is less effective than strain \(C\) by generating energy from a given amount of resource, unless a very large amount of resource is gathered to be converted into energy. This is ensured when the ratio \(\Gamma_{\text{ATP}} = A_D^{\text{ATP}}/A_C^{\text{ATP}}\) is larger than 1. This allows a fast metabolic pathway for \(D\) strains to produce more energy in the case where they capture a much higher amount of resource than the efficient strain. This possibility is particularly important as it is known that many cells can work as respiro-fermenting cells, concomitantly using the two alternative pathways of ATP production. Such respiro-fermentative metabolism is a typical mode of ATP production in unicellular eukaryotes like yeast [6,16]. In such a situation, the yield of the whole process is reduced as the cell drives more resource to be metabolized inefficiently, a habitual behaviour under the condition of plentiful resource. Of course, the situation \(A_D^{\text{ATP}} < A_C^{\text{ATP}}\) can also be seen, but the opposite \(A_D^{\text{ATP}} > A_C^{\text{ATP}}\) represents the worst scenario for the efficient strain. If it can thrive in such a case, it will naturally be favoured under less harsh conditions.

2.3. Social dilemma

Besides setting down the constraints on the biologically relevant parameter values, it is also important to find the region of the parameter space at which the rate–yield trade-off in fact produces a tragedy of the commons [1]. As commonly defined, the social dilemma exists if in a pairwise competition for a common resource strain \(D\) outcompetes strain \(C\) [2,5,29].

For a given amount of resource \(S^*\), the uptakes of resource by strains \(C\) and \(D\) are, respectively,

\[
S_C = \frac{\dot{J}_C^S}{\dot{J}_C^S + \dot{J}_D^S} S^* \quad (2.5)
\]

and

\[
S_D = \frac{\dot{J}_D^S}{\dot{J}_C^S + \dot{J}_D^S} S^*. \quad (2.6)
\]

The ratio between \(J_D^{\text{ATP}}\) and \(J_C^{\text{ATP}}\) provides us the relative advantage of strain \(D\) over strain \(C\) in a pairwise competition. From equations (2.4), we obtain that

\[
r = \frac{J_D^{\text{ATP}}}{J_C^{\text{ATP}}} = \frac{\Gamma_{\text{ATP}}}{\Gamma_{\text{ATP}}} \frac{1 - \exp(-a_D^{\text{ATP}}S_D)}{1 - \exp(-a_C^{\text{ATP}}S_C)} \quad (2.7)
\]

If this ratio is larger than 1, the social dilemma holds. This reasoning will be employed to map the set of parameter values where the social dilemma exists. As the ratio \(r\) depends on the amount of resource,
in order to delimit the social dilemma region in a population of arbitrary size, we must use $S^* \simeq S/N$, instead of using $S$, the total amount of resource. $N$ denotes the population size. As observed from the simulations, the system evolves in such a way that $S/N$ is always small at equilibrium. As we will see later, the limit of small $S/N$ is useful to understand the results.

3. Results

We will start this section by presenting some analytical approximations for well-mixed populations. These calculations will later serve as a guidance to determine the situations in which simulation results match the theoretical prediction. They also help us to delimit the situations where phenomena driven by stochastic effects manifest, which are not captured by the analytical approach carried out here. As in the analytical approximations no population structure is assumed, the results will be particularly instructive to understand whether group structure is a driving force for the maintenance of the efficient strain.

3.1. Analytical results for well-mixed populations

3.1.1. Equilibrium size of a population of strains with fast metabolism

Let us propose a discrete-time model for a well-mixed population with only cells of type $D$. More generally, the population size in time step $t + 1$ can be written as

$$n(t + 1) = n(t) + g(n(t), S)n(t) - \nu n(t),$$

where the function $g(n(t), S)$ denotes the growth rate of strain $D$, which depends on the population size as well as on the availability of resource per time step, $S$. The death rate is denoted by $\nu$ (as a mathematical guide for the analytical development to be done, see [30]). In the above equation, it is implicitly assumed that $g(n(t), S)$ is proportional to the amount of resource captured by the consumer. The conversion of resource into energy (directly translated as growth rate) occurs into two steps: uptake of resource and its posterior conversion into energy. As there is only one type of strain in a homogeneous population, the amount of resource per individual at time $t$ is just $S/n(t)$. Following equation (2.4d), the growth rate is equivalent to $g(n(t), S) = aD[1 - \exp(-aD^\text{ATP} S/n(t))]$, where $a_D = A_D^\text{ATP}/E_{\text{max}}$, which is a necessary rescaling because individuals only replicate if their internal energies surpass the threshold $E_{\text{max}}$. Plugging this expression into equation (3.1) we get

$$n(t + 1) = n(t) + aD \left[1 - \exp\left(-\frac{aD^\text{ATP} S}{n(t)}\right)\right] n(t) - \nu n(t).$$

The system is under equilibrium when $n(t + 1) = n(t) = \hat{n}$ and has two equilibrium solutions, $\hat{n} = 0$ and $\hat{n} = -aD^\text{ATP} S/\ln(1 - \nu/aD)$. Note that the second solution is only meaningful if $\nu < a_D$.

3.1.2. Stability analysis

Next, we analyse the local stability of the two equilibrium solutions. By writing the dynamics of the discrete-time model as $n(t + 1) = f(n(t))$, a solution, $n = \hat{n}$, is said to be stable when $-1 < \lambda < 1$, where $\lambda = (df/dn)|_{n=\hat{n}}$. Note that the derivative of $f(n(t))$ with respect to $n$ must be evaluated at the equilibrium. A value of $|\lambda| < 1$ warrants that any small perturbation from the equilibrium solution will be damped, bringing the system back to equilibrium. Whenever $|\lambda| > 1$, any disturbance drives the system away from the equilibrium point.

For the first equilibrium, $\hat{n} = 0$, $\lambda$ is evaluated as $\lambda = (df/dn)|_{n=0} = 1 - \nu + aD$, and so the solution is stable when $\nu > aD$, i.e. $\nu E_{\text{max}} > A_D^\text{ATP}$. On the other hand, the second equilibrium, $\hat{n} = -aD^\text{ATP} S/\ln(1 - \nu/aD)$, provides $\lambda = (df/dn)|_{n=-aD^\text{ATP} S/\ln(1 - \nu/aD)} = 1 + aD(1 - \nu/aD) \ln(1 - \nu/aD)$. Because $\nu < a_D$ in the range of validity of the solution, the logarithm is always negative, meaning $\lambda < 1$. Therefore, the second solution is always stable in the range it is valid.

3.2. Evolutionary invasion analysis

3.2.1. Invasion by the efficient strain

A further step of our analytical calculation is to check under which conditions a population of strain with rapid growth can be invaded by a mutant that uses the resource more efficiently. The simulations start from an isogenic population of strain $D$. After the population reaches the stationary state, a single
type C individual is introduced. The number of individuals of strains D and C over time, \( n_D \) and \( n_C \), are determined by the following set of equations:

\[
n_D(t + 1) = n_D(t)[1 + a_D[1 - e^{-a_D^{ATP}/S/\left(1 - n_D + n_C\right)}] - \nu]
\]

and

\[
n_C(t + 1) = n_C(t)[1 + a_C[1 - e^{-a_C^{ATP}/S/\left(1 - n_D + n_C\right)}] - \nu],
\]

where \( a_C = A_C^{ATP}/E_{max} \). It is important to point out that as there are two distinct strains, the amount of resource captured per individual is no longer \( S/n(t) \) but obeys the resource partition as calculated by equations (2.1) and (2.2). The dynamics is driven by a density-dependent growth rate. One possible equilibrium of the above system is

\[
\hat{n}_D = -\frac{\alpha_D^{ATP}S}{\ln(1 - \nu/a_D)} \quad \text{and} \quad \hat{n}_C = 0.
\]

This equilibrium corresponds to the non-zero equilibrium of the resident population of strain D, as shown in §3.1.1. As aforementioned, the system is self-organized in such a way that the population size is reflected by the supply of resource into the system and also determined by the metabolic properties of its individuals. A less efficient metabolism entails that the accumulation of internal energy is slower. In such case the system cannot sustain a large population size (electronic supplementary material, figure S1). It is worth mentioning that as the regime of small \( \nu/a_D \) is considered, it ensues that taking this limit into equation (3.5) the per capita resource availability \( S/n \) is proportional to \( \nu/a_D \) and hence also pretty small. This finding is always observed in our simulations. Next, we analyse the stability conditions of this equilibrium under the invasion of a small quantity of efficient strains. For multiple-variable models, the local stability of the equilibrium solutions is found by looking at the eigenvalues of the Jacobian matrix of the system.

After determining the eigenvalues and setting the conditions for stability (see the electronic supplementary material), it is possible to determine a relationship delimiting the parameter region at which the efficient strain can invade the resident population of type D cells. For fixed \( A_C^{ATP} \) and variable \( A_D^{ATP} \), we find that

\[
\Gamma_{ATP} = \frac{\nu/a_C}{1 - (1 - \nu/a_C)\Delta^{ATP}\epsilon},
\]

where \( \Delta^{ATP} = \alpha_D^{ATP}/\alpha_C^{ATP} \) and \( \epsilon = A_D/A_C \), as previously defined. Equation (3.6) is used for the comparison with the isocline \( P_{fix} = 1 \), the relative fixation probability equal to 1, obtained from simulations. Therefore, our finding shows that the solution, equation (3.5), is stable against invasion above the curve (3.6), while the strain C can succeed to invade in the parameter space under the curve.

3.2.2. Invasion by the inefficient strain

Now the opposite situation is considered—the conditions under which a mutant with rapid metabolism (type D) can invade a population dominated by cells of type C. The calculations are straightforward as calculations are similar as in §3.2.1. As the resident population is initially composed of cells type C only, its size at equilibrium is \( n_C = \hat{n}_C = -\alpha_C^{ATP}S/\ln(1 - \nu/a_C) \), and \( a_C = A_C^{ATP}/E_{max} \).

The Jacobian matrix is now evaluated at

\[
\hat{n}_C = -\frac{\alpha_C^{ATP}S}{\ln(1 - \nu/a_C)} \quad \text{and} \quad \hat{n}_D = 0
\]

as the strain D is assumed to be rare. Likewise, a relation between \( \Gamma_{ATP} \) and \( \Delta^{ATP} \) is obtained, and the curve now delimits the parameter region at which the efficient strain is stable against invasion by the inefficient strain. It is found that

\[
\Gamma_{ATP} = \frac{\nu/a_C}{1 - (1 - \nu/a_C)\Delta^{ATP}\epsilon},
\]

which is exactly equal to equation (3.6). However, now the solution, equation (3.7), is stable under the curve. Thus, we conclude that the parameter region \( \Gamma_{ATP} \) versus \( \Delta^{ATP} \), where the efficient strain is evolutionarily stable against invasion by the selfish strain D, coincides with the region at which it can, while rare, invade an isogenic population of type D individuals.
but it evidences a selective advantage for strain \( C \) \( \text{Saccharomyces cerevisiae} \) empirically measured in populations of individuals of type \( r \) \( (r = 1 \text{ in equation (2.7). This line overlaps the lines given in equations (3.6) and (3.8). The grey region denotes that the defector population goes extinct before the cooperator is introduced. The other parameter values are resource amount \( S = 25 \), death rate \( \nu = 0.01 \), group carrying capacity \( P_{\text{max}} = 10 \), internal energy for splitting \( \epsilon = 0.1 \text{ the efficient strain is already unlikely to invade and to fixate.} \)

3.3. Simulation results

In this section, we present the simulation results of the evolutionary invasion analysis. In the electronic supplementary material, we also provide simulation results and plots of the population size of pure \( D \) strains at equilibrium. This is critical as the population size determines the strength of stochasticity on the dynamics of the system.

3.3.1. The relative fixation probability of a single individual of the efficient strain

From the heat map in figure 2, we observe how the relative fixation probability is influenced by the ratios \( \Gamma_{\text{ATP}} = \frac{A_D}{A_C} \text{ATP} \) and \( \Delta_{\text{ATP}} = \frac{\alpha_D}{\alpha_C} \text{ATP} \). The ratio \( \epsilon = \frac{A_D}{A_C} \) is set at 10, which is a typical value, as empirically measured in populations of \( \text{Saccharomyces cerevisiae} \) \[6\]. Making the comparison with well-mixed populations, over a broader domain of the parameter space the efficient strain is selected for in the structured populations. Even when \( \Delta_{\text{ATP}} \) is already close to 1, meaning that the yield of strain \( D \) is almost equivalent to strain \( C \), the efficient strain can still persist. Overall, the domain at which the efficient strain is favoured in well-mixed populations is considerably restricted in comparison to structured populations. For instance, around \( \Delta_{\text{ATP}} = 0.1 \) the efficient strain is already unlikely to invade and to fixate.

The curves corresponding to equations (3.6) and (3.8) are represented by the green line in figure 2a,b. The region below the green line portrays the domain at which the efficient strain can invade and replace a population of individuals of type \( D \). Of course, the outcome of the dynamics is not deterministic, but it evidences a selective advantage for strain \( C \). The white line is the isocline \( P_{\text{fix}} = 1 \) (the relative fixation probability is equal to 1). The line setting out the onset of the social dilemma corresponding to \( r = 1 \text{ in equation (2.7). This line overlaps with the green line, equations (3.6) and (3.8), and thence it is not perceivable in the plot. Indeed, in the limit \( S/N \) small, which always holds when \( \nu/\alpha_D \) is also small (see equation (3.5)), the two curves are essentially the same and become \( \Gamma_{\text{ATP}} = 1/\Delta_{\text{ATP}} \). This observation leads us to conclude that in well-mixed populations the efficient strain is an effective intruder only in the domain where there is no dilemma. Note that the social dilemma line becomes independent of \( S \) in the limit \( S/N \) small.

On the other hand, in structured populations the efficient strain is selectively advantageous over a much larger set of the parameter space. The realm of the high yield strain goes further beyond the line setting out the social dilemma. The isocline \( P_{\text{fix}} = 1 \) is considerably shifted to the right. The region
Figure 3. Relative fixation probability as a function of $\Delta_{\text{ATP}}$ for fixed values of $\Gamma_{\text{ATP}}$. The blue triangles show the simulation outcomes for homogeneous populations, whereas the red dots for structured populations. The data points correspond to 40 distinct populations and for each population 100 000 independent runs were performed. The black horizontal line denotes relative fixation probability 1. The parameter values are $S = 25$, $\nu = 0.01$, $P_{\text{max}} = 10$, $E_{\text{max}} = 10$, $A_D = 10$ and $A_C = 1$.

between the social dilemma line (green line) and the isocline $P_{\text{fix}} = 1$ is of special interest, as it shows the prominent role of structuring in promoting the fixation and maintenance of the strain exploiting the resource efficiently.

Looking at the heat map, it is notable that there exists a non-straightforward relationship between $\Delta_{\text{ATP}}$ and $\Gamma_{\text{ATP}}$. At fixed $\Gamma_{\text{ATP}}$ the relative fixation probability approaches a two-humped function of $\Delta_{\text{ATP}}$, being maximized at low and intermediate values of $\Delta_{\text{ATP}}$ (figure 3). In each of the panels of figure 3, $\Gamma_{\text{ATP}}$ is kept constant while $\Delta_{\text{ATP}}$ varies. Well-mixed populations (blue triangles) display a smoother scenario and the relative fixation probability is now a one-humped function of $\Delta_{\text{ATP}}$. The point at which the relative probability is maximized depends on $\Gamma_{\text{ATP}}$, being shifted towards higher values of the ratio $\Delta_{\text{ATP}}$ as $\Gamma_{\text{ATP}}$ decreases. In well-mixed populations the selective advantage of strain C over strain D is confined to the region where there is no dilemma.

Yet, in structured populations the range of $\Delta_{\text{ATP}}$ in the plot embodies three distinct regions: the region at which there is no dilemma, a second region where the dilemma exists and the efficient strain has a selective advantage, and a third one at which the dilemma still holds but the efficient strain is no longer able to persist. The first peak occurs at small $\Delta_{\text{ATP}}$ as observed in well-mixed populations, but now there is a second peak that lies at the range of intermediate values of $\Delta_{\text{ATP}}$. The first peak is easily understood as it lies in the region of no social dilemma and the efficient strain already has a selective advantage. In this region, the ratio $r$, as defined in equation (2.7), is greater than 1. In fact, the advantage of the efficient strain over the inefficient one is observed in both structured and unstructured populations, emphasizing that cooperators are always favoured in this regime.

The second peak is more difficult to analyse, as we have the coupling between the dynamics of group formation and the within-group dynamics. At a first look, it may seem counterintuitive that the relative fixation probability increases with defector efficiency at intermediate values of $\Delta_{\text{ATP}}$. For explaining this result, we need to take into account how both local competition for resource and group expansion are affected. On the one hand, a more efficient defector (larger $\Delta_{\text{ATP}}$) leads to an increased strength of local competition for resource as experienced by the cooperator; on the other hand, it also allows a faster expansion of the group containing the invading cooperator, thus favouring group division and the spreading of the cooperative trait. Thus, the net advantage of the efficient strain comes from the net outcome of these two competing mechanisms. At very large values of $\Delta_{\text{ATP}}$, the local competition as experienced by the efficient strain is so strong, in such way it cannot reproduce inside the group, and it outweighs the advantage brought about by the group expansion.

Next, we study how the relative fixation probability changes with the group carrying capacity, $P_{\text{max}}$. This enables one to assess the strength of stochasticity on the population dynamics. Previous studies have claimed that in multilevel selection the likelihood of invasion and posterior fixation of cooperative
traits are particularly enhanced for small group sizes [23], evidencing the importance of stochasticity in the dynamics [31]. Similar behaviour was observed in a resource-based model of pairwise interactions within the multilevel selection framework [32]. The contour map in figure 4a displays the relative fixation probability for different values of \( P_{\text{max}} \) and different values of \( \Delta_{\text{ATP}} \). For extremely small group sizes, like \( P_{\text{max}} = 5 \), the relative fixation probability varies enormously with the ratio \( \Delta_{\text{ATP}} \), being maximized at intermediate values. In such small groups, the relative fixation probability can be 60-fold higher than the relative fixation probability of a neutral trait. This occurs because as smaller groups are considered there exists a greater chance of formation of uniform groups (comprised only cooperators) during the formation of new groups at an early stage of the introduction of the cooperative trait. Once a uniform group is produced, the cooperators are free of the exploitation of the defectors and the group enjoys a higher chance of fixation. At large values of \( P_{\text{max}} \), instead of a pronounced change with \( \Delta_{\text{ATP}} \), the probability is roughly constant at a broad interval range of \( \Delta_{\text{ATP}} \), as inferred from the isoclines. Of particular interest is the isocline delimiting the onset of the phase at which the efficient strain is favoured. From the isocline \( P_{\text{fix}} = 1 \), one deduces that for \( \Delta_{\text{ATP}} \) between 0.1 and 0.4 there is a well-defined critical value for the group carrying capacity, \( P_{\text{max},c} \approx 17 \), beyond which the efficient strain becomes counter-selected. Although as the efficiency of the strain \( D \) is further reduced (small \( \Delta_{\text{ATP}} \)), the critical group size \( P_{\text{max},c} \) increases again. For the sake of completeness, figure 4b shows that the relative fixation probability is a well-behaved monotonic decreasing function of \( P_{\text{max}} \) under fixed \( \Delta_{\text{ATP}} \).

3.4. Enabling migration between groups

As aforementioned, increasing the group carrying capacity \( P_{\text{max}} \) will reduce the strength of stochasticity and thus creating a less favourable environment for the promotion of the efficient usage of resource.
Figure 5. The effect of the consumption rate. Heat map of the relative fixation probability in terms of the ratio $\Gamma_{\text{ATP}}$ and the consumption rate of defectors $A_D$. The parameter values are $S = 25$, $\nu = 0.01$, $P_{\text{max}} = 10$, $E_{\text{max}} = 10$ and $\Delta_{\text{ATP}} = 0.2$. The thick white lines correspond to isolines. The data points correspond to an average over 10 distinct populations and for each population 10 000 independent runs were performed.

Figure 6. Effect of migration on the relative fixation probability. The relative fixation probability is plotted as a function of the migration rate $m$ for three distinct values of carrying capacity $P_{\text{max}}$. The parameter values are $S = 25$, $\nu = 0.01$, $E_{\text{max}} = 10$, $A_D = 10$, $A_c = 1$, $\Gamma_{\text{ATP}} = 1.5$ and $\Delta_{\text{ATP}} = 0.5$ (blue points) and $\Delta_{\text{ATP}} = 0.04$ (red points).

Here, we present the effects of an alternative natural force, migration between groups. At each time step every individual has a probability $m$ to leave its original group and move to a randomly chosen group in the population. This is known in ecological studies as the island model of stochastic migration [33]. By allowing migration to happen, the effective population size can change and in the limit of high migration rates it is expected that the structured population behaves similar to its well-mixed counterpart.

The effect of the migration rate on the relative fixation probability can be seen in figure 6. Different situations are considered in the face of our previous knowledge about the system. Distinct values of $\Delta_{\text{ATP}}$ and of carrying capacity $P_{\text{max}}$ are considered. At $\Delta_{\text{ATP}} = 0.5$ (blue points) the efficient strain is selected against in well-mixed populations but not in structured populations. At $\Delta_{\text{ATP}} = 0.04$ (red points), the efficient strain is favoured in both scenarios. In the latter situation, migration does not affect the outcome and the relative fixation probability is invariant.

Important information about the role of migration is gathered from the analysis of the case $\Delta_{\text{ATP}} = 0.5$. Starting from small migration rates, which means the groups are more isolated, the relative fixation probability remains roughly constant over a wide range of migration rates, and then sharply goes to zero when the migration rate increases further, signifying less isolation between groups. The first plateau coincides with the value of the relative fixation probability in our structured population model, while
after the transition the situation resembles well-mixed populations at which the efficient strain can no longer thrive.

4. Discussion

We have studied the evolutionary dynamics of a population with two strains under a trade-off between rate and yield. The two strains with distinct metabolic pathways compete under a scenario of limited resource. This problem has long been debated within the framework of evolution of cooperation, as the strain that uses resource efficiently is described as a cooperator, while the strain with high uptake rate, at the expense and degradation of its own environment, is said to display a defecting behaviour [5,34]. The trade-off between high rate and yield occurs due to biophysical restraints. Experimental results show a clear negative correlation between resource uptake and efficiency of conversion of resource into energy [3,5,35]. The existence of such trade-offs in heterotrophic microorganisms, like bacteria and yeast, and the fact that the efficient manipulation of resource was made possible with the appearance of the metabolic pathway that makes use of oxygen, the so-called respirers, have motivated a large body of studies. A clear motivation is the issue related to the major transition from single cells towards multicellular forms of life [36,37]. Multicellularity can thus be thought of as the integration of smaller units to build up a more complex machinery.

Here we proposed an explicit resource-based model in which competing metabolic strategies are assumed. We studied the conditions under which the cooperative strain with high yield metabolism can invade and outcompete a resident population of selfish strains. By investigating structured and well-mixed populations, we understand better whether the existence of group structure is a required condition for the maintenance of efficient metabolism pathways.

Our results show that in well-mixed populations, the efficient strain is only viable in a subset of the parameter space corresponding to the region where there is no social dilemma. In other words, the efficient strain can only invade a population of inefficient strains when it is rare, under the condition it is already a strong competitor in a pairwise competition with its opponent. The lines delimiting the social dilemma region and the isocline $P_{fix} = 1$ overlap. These lines also delimit the region of stability of the two competing metabolic strategies. Together, these results suggest that a complete mixing of cells creates an unfavourable scenario for the appearance and subsequent maintenance of an efficient mode of metabolism.

On the other hand, in structured populations a more favourable scenario is presented. Over a broad domain of the parameter space, which goes further beyond the boundaries set down by the social dilemma and the invasion analysis, the efficient strain is selected for. This result highlights the role of structuring, also corroborated by the study of migration effects on the dynamics of fixation of the efficient strain. As migration is reduced and so isolation increases, an abrupt augmentation of the relative fixation probability of the efficient strain is observed. Although local competition among relatives may increase in structured populations [38,39], it does not mitigate the benefits of increased stochasticity in the model. More evidence of this finding is obtained by checking the dependence of the relative fixation probability on the carrying capacity $P_{max}$, whereby it is demonstrated that the efficient strain faces an even more favourable environment when smaller groups are assumed.

One question remains: which mechanisms can relieve the effect of increased competition in small groups? Local competition for resource is attenuated in the structured populations as groups are elastic and their sizes can vary between fission events. This is consistent with previous findings [39,40]. Although a larger group means less resource available per individual, there is no competition for space. Once one individual reproduces it does not necessarily lead to the removal of any interactant. Additionally, a group containing more efficient individuals has as advantage, as a larger amount of resource will be available and the net amount of energy converted in the process can be larger. This allows the group to reach the carrying capacity $P_{max}$ faster. The third mechanism is the process of group split, which produces a twofold benefit: first by reducing local competition because after division the number of cells per group drops; second, if the number of efficient cells that occupy a given group increases, so does the chance that a newborn group is only filled by efficient strains, thus helping to perpetuate the trait. In summary, three premises of the model—group expansion, overlapping generations and group split—contribute to attenuate local competition.

The observation of a critical value of $P_{max}$ imposes a threshold for the level of finiteness to act effectively. The group split is a key premise of the model. In a recent work, Ratcliffe et al. [19] designed an experiment to understand the evolution of complex multicellular organisms from unicellular ancestors.
In the experiment, a population of yeast (*S. cerevisiae*) was artificially selected for multicellularity by growing in nutrient-rich medium. Serial transfers of the bottom fraction of the liquid were performed, and cells collected together with the liquid formed the next generation. After a few transfers, they already observed that populations were dominated by snowflake-shaped multicellular clusters. Their witness of the occurrence of cluster reproduction is a basic premise to settle down life in its multicellular form. The authors ascertained that the daughter clusters were produced as multicellular propagules rather than mass dissolution of parental clusters. They also observed the existence of a threshold for the minimum size of daughter clusters [19]. In our viewpoint, the findings in the experiment of Ratchiff et al. are in line with the assumptions of the group dynamics of the proposed model in this work.

As conjectured, the first step in the evolutionary transition from single-celled to multicellular organisms most likely involved the evolution of undifferentiated multicellularity [36]. Our results show that it is very unlikely that the efficient mode of ATP production becomes widespread in a scenario of global interaction and competition (well-mixed populations). Even though efficient strains can arise through mutation, the maintenance of those strains in a well-mixed population would be very improbable as they could be easily invaded by strains with higher growth rate in spite of the low efficient metabolism. Alternatively, the appearance and subsequent growth of efficient strains in frequency must come together with the emergence of a mechanism that prevents dividing cells from complete separation enabling clusters through parent–offspring adhesion [19]. Collecting efficient feeders into a group, whether they are of the same kin or not, would benefit them equally well.

Data accessibility. Our C++ codes are available at Dryad Repository in http://dx.doi.org/10.5061/dryad.q6784 [41].

Authors’ contributions. A.A. conceived the model, performed the simulations, developed the analytical calculations and drafted the manuscript. L.F. participated in the model discussion, performed the simulations, developed the analytical calculations and critically revised the manuscript. W.H. participated in the model discussion and drafted the manuscript. F.F.F. conceived the model, developed the analytical calculations and drafted the manuscript. P.R.A.C. designed the study, conceived the model, developed the analytical calculations and drafted the manuscript. All authors gave final approval for publication.

Competing interests. We have no competing interests.

Funding. A.A. has a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). L.F. has a fellowship from the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE). W.H. gratefully acknowledges funding from the Max Planck Society. F.F.F. is supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). P.R.A.C. is partially supported by the Brazilian research agency CNPq and acknowledges financial support from the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) under project no. APQ-0464-1.05/15.

References


