Metabolically Coupled Replicator Systems: Overview of an RNA-World model concept of prebiotic evolution on mineral surfaces

Tamás Czárán[1, *]
[1]MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, H-1117 Pázmány Péter sétány 1/c, Budapest, Hungary
Email: czaran@caesar.elte.hu

Balázs Könnyű [2]
[2]Eötvös Lorand University, Department of Plant Systematics, Ecology and Theoretical Biology, H-1117 Pázmány Péter sétány 1/c, Budapest, Hungary
Email: konnyu@caesar.elte.hu

Eörs Szathmáry[1, 2, 3]
[1]MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, H-1117 Pázmány Péter sétány 1/c, Budapest, Hungary
[2]Eötvös Lorand University, Department of Plant Systematics, Ecology and Theoretical Biology, H-1117 Pázmány Péter sétány 1/c, Budapest, Hungary
[3]Parmenides Institute for the Conceptual Foundations of Science, Kirchplatz 1, D-82049 Munich/Pullach, Germany
E-mail: szathmary.eors@gmail.com

*Corresponding author

Abstract

Metabolically Coupled Replicator Systems (MCRS) are a family of models implementing a simple, physico-chemically and ecologically feasible scenario for the first steps of chemical evolution towards life. The hypothetical starting point of the scenario is a large population of RNA(-like) macromolecules produced abiotically on a suitable spot of prebiotic Earth, attached to a mineral surface capable of binding both the macromolecules and the monomers they are made of. Evolution sets in as soon as any one of the RNA molecules become *autocatalytic* by engaging in template directed self-replication from activated monomers, so that its population size starts increasing exponentially. Competition for the finite external supply of monomers ignites selection favouring RNA molecules with catalytic activity helping self-replication by any possible means. The most straightforward way of providing such catalytic help is to become a *replicase ribozyme* offering a new self-copying mechanism, even if it is only marginally more efficient than the one available before. An additional way is through increasing monomer supply by contributing to monomer synthesis from external resources, i.e., by evolving *metabolic enzyme activity*. *Retroevolution* may build up an increasingly autotrophic, cooperating community of metabolic ribozymes running an increasingly complicated and ever more efficient metabolism.

Maintaining such a cooperating community of metabolic replicators raises two serious ecological problems: one is keeping the system *coexistent* in spite of the different replicabilities of the cooperating replicators; the other is *constraining parasitism*, i.e., keeping "cheaters" in check. Surface-bound MCRS provide an automatic solution to both problems: the coexistence of cooperating replicators and their parasite resistance are the consequences of assuming the local nature of metabolic interactions. In this review we present an overview of results published in previous articles, showing that these effects are, indeed, robust in different MCRS implementations, by considering different environmental setups and realistic chemical details in a few different models. We argue that the MCRS model framework naturally offers a suitable starting point for the future modelling of membrane evolution and extending the theory to cover the emergence of the first protocell in a self-consistent manner. The coevolution of metabolic, genetic and membrane functions is hypothesized to follow the *progressive sequestration* scenario, the conceptual blueprint for the earliest steps of protocell evolution.

Keywords: early molecular community, stability, coexistence, spatially explicit model, cellular automata

Graphical Abstract

1. Introduction

The problem of the origin of life is a scientific question, but one with a strong historical dimension. The historical aspect raises at least two difficulties which seem impossible to overcome. First, those who venture into the field of prebiotic evolution should be prepared to accept the fact that very likely noone will ever be able to factually verify or falsify claims on any hypothetical series of events that would have produced the first living organism, simply because no fossil proof of any kind can be hoped for from the enormous distance of over 3 billion years ago to support such hypotheses. Second, we have no clue on what alternative histories of prebiotic chemical evolution could have existed on the prebiotic Earth, since the chemical universalities of all recent organisms suggest that the actual history is unique, although it is well possible that different attempts had been made by radically different prebiotic chemical systems, and the one that successfully launched life as we know it today had won the competition between those possible candidate systems at a very early phase.

Therefore, studying the process of prebiotic evolution is largely restricted to the domain of

the *possible*, not the *actual*: we may search for scenarios that are feasible from a physical-chemical point of view and are reconcilable with the chemical organization of recent forms of life. Even though we cannot tell with certainty what actually happened, we may have reasonably strong scientific arguments to decide what could have happened and what not (Eschenmoser, 2007). *Systems chemistry* (von Kiedrowski et al. 2010) offers a wide range of theoretical and experimental methods for constructing and testing possible evolutionary scenarios of prebiotic evolution, from the very beginning to the emergence of the first living cell. We attempt to sketch such a scenario in this paper, one that we believe is both feasible and open to further improvements through the inclusion of more detailed and more realistic physical and chemical mechanisms.

Molecular interactions had shaped the chemical evolution of prebiotic macromolecular structures with a selective power almost as efficient as that of competitive and sexual interactions driving the evolution of living creatures today. In view of all that we know – or suspect – about the earliest phases of the origins of life these molecular interactions may have played the key roles in the transformation of matter from inanimate to animate.

Most students of the origin of life agree that chemical evolution must have started with the formation of small organic molecules (like formaldehyde, hydrogen-cyanide etc.) through abiotic (geo)chemical processes. Lacking sufficiently accurate information of the climatic and geochemical environment on Earth four billion years ago, the study of even this initial step of the wake of life is largely speculative and often controversial with respect to the actual details (Martin and Russel 2003; Miller, 1953; Miyakawa et al. 2002; Monnard et al. 2003; Powner et al. 2009; Wächtershäuser, 1990), so much so that the hypothetical initial sets of prebiotic organic compounds show a large variety across the literature. Whatever their chemical identities were, those small organic molecules must have reacted with each other to produce macromolecules which later formed macromolecular complexes or communities by hypothesized self-assembly processes or coexistence mechanisms of different kinds (Chen and Walde, 2010; Cleaves et al. 2012; Deamer and Weber, 2010; Ehrenfreund and Cami, 2010; Ferris, 2006; Garay, 2011; Johnson et al. 2008; Miller, 1953; Miyakawa et al. 2002; Orgell, 2004; Powner et al. 2009; Rushdi and Simoneit, 2001). The mechanisms of self-assembly and macromolecular community formation are often theoretically problematic, either because the assumptions of the underlying (toy) models are too schematic or because they are physically or chemically unrealistic (Morowitz et al. 2000; Pross, 2004; Szathmáry, 2006; Segré et al. 2001). The actual chemical and evolutionary details of the many different scenarios are usually implicit, so it is often difficult to see how the envisioned macromolecular complex or community could be a self-sustaining and self-regulated unit of life or of evolution (Gánti, 1987; Rasmussen et al. 2009).

Historically the first prebiotic replicator community model was Eigen's *hypercycle* (Eigen and Schuster, 1979). It was conceived to solve the chicken-and-egg problem of reliable replication: one would need a long replicase in the first place that would be able to accurately copy itself. This poses the question how the first such long replicase could have emerged and persisted without the copying accuracy required? Lacking an efficient replicase the copying process is hampered by frequent errors, leading to an *error catastrophe* (Eigen and Schuster, 1979) for sequences longer than the *error threshold*. The hypercycle was thought to solve the problem by splitting the long sequence into short ones which are not prone to the error catastrophe. To avoid competition among the fragments they are organized in a structure such that they help each other's replication in a cyclical topology. The cumulated size of the members in a hypercycle may exceed the size limit set by the error threshold for single sequences. Theoretical considerations have proven that the simple hypercycle cannot be evolutionarily stable (Boerlijst, 2000; Boerlijst and Hogeweg, 1991; Bresch et al. 1980; Kim and Jeong, 2005). Two types of mutants both may ruin the cooperation of the replicators in a hypercycle: *selfish parasites* (replicators helping themselves but not the downstream neighbour in the cycle) cut the cyclical flow of benefits, whereas *shortcut parasites* (helping another member of the hypercycle instead of the downstream neighbour) exclude some members

from the circular flow of benefits, repeated shortcut mutations ultimately reducing the hypercycle to a single member (Fig.1.).

1.1. The Metabolic Replicator paradigm

Our approach, like the hypercycle, is one of the "evolution of coexistent replicator communities" scenarios, aimed at explaining the dynamical stability and the evolvability of a hypothetical community of macromolecules provided by an initially random RNA World (Gilbert, 1986; Joyce, 2002). The *Metabolically Coupled Replicator System (MCRS)*, we believe, is the most feasible candidate suggested so far for a prebiotic chemical supersystem that may have evolved into the first cellular form of life, the *protocell*. We shall explain below why we think so.

The central proposal of the RNA-World scenario is that the first evolvable entities on prebiotic Earth may have been RNA (or RNA-like – (Eschenmoser, 2007; Hall, 2004; Robertson and Joyce, 2010)) macromolecules, and the first protocell is the evolutionary product of an RNA(-like) macromolecular community, the members of which were connected by a specific set of mutually advantageous interactions. This suggestion is appealing for a number of very important reasons. RNA inherently embodies the first two of the three indispensable *infrabiological* (Szathmáry et al, 2005) components of living systems (*metabolism*: catalytically channelled reaction network producing compounds necessary for reproduction; *genetics*: hereditary information transmission through template replication; and *membrane*: partial separation of the biological entity from the outside world). First, RNA has been proven to possess a wide range of enzymatic activities (Chen et al. 2007; Landweber et al. 1998; Lilley, 2003) that are absolutely necessary for driving even a very primitive metabolism. Second, RNA is inherently modular, i.e., it is composed of a few chemical modules (nucleotides) in a linear arrangement, so that the sequence of the modules may carry information. Sequences may be of a virtually unrestricted variety, and the unambiguous complementation of the modules allows for the template replication of the sequences (Szathmáry, 2006), i.e., for genetic information transmission through generations of RNA molecules. It is this inherent dual (metabolic and genetic) role of RNA which earned the name "*Metabolically Coupled Replicator System*" to our prebiotic evolutionary scenario. The third infrabiological component (membrane) may be the product of subsequent evolution within the metabolic RNA community, or its function might have been initially supplied by specific environmental conditions, as we will show later (Branciamore et al. 2009).

Of course there are large gaps in our knowledge with respect to prebiotic chemistries capable of delivering activated modules for the replication of early RNA-World molecules. Yet, we have no other choice at present but assuming that the RNA-World was initially absolutely "heterotrophic", that is, the first RNA(-like) macromolecules were randomly assembled from activated modules which in turn were the products of so far largely unknown geochemical processes. "Black smokers" (hot and high pressure volcanic vents thousands of meters below sea level in the ocean-beds) seem to be reasonably good candidates for having supplied the modules (Deamer and Weber, 2010; LaRowe and Regnier, 2008; Orgel, 2004), but we are still far from even an established hypothesis on this topic. Fortunately, there is much more known about the possibilities of non-template-directed RNA synthesis from activated monomers (nucleotides). Experimental results suggest that mineral (clay) surfaces – like that of montmorillonite – can catalyse spontaneous bond formation between activated nucleotides (Ferris, 2006, Ferris et al. 1999), resulting in RNA molecules of different lengths and random nucleotide sequences.

Once a sufficiently diverse random set of RNA molecules is available, the stage is set for the evolution of a sustainable, cooperative RNA-World scenery to play out (Copley et al. 2007; Manrubia and Briones, 2007). An essential criterion of this to happen is that the RNA molecules originally produced by spontaneous bond formation become *template replicated*. In fact this is the single most crucial condition for evolution to set in and select for RNA assemblies somewhat more

efficient in replication than others. It seems very reasonable to assume that some random RNA molecules – or an assembly of a few different ones – generated on the mineral surface might have had a weak RNA-replicase activity. This would have been sufficient to ignite the selection process for a gradual increase of replicase activity (Attwater et al. 2010; Johnson et al. 2001). The snag with this straightforward reasoning is empirical: even though RNA replicase ribozymes are actively searched for in many laboratories (Johnson et al. 2001), the ones discovered so far are incapable of replicating themselves because they are longer than the longest template they can handle (Attwater et al. 2010; Wochner et al. 2011). In spite of the lack of a real breakthrough in this respect so far, this is one of the most promising directions of experimental research on prebiotic evolution: we seem to be quite close to having a proper RNA replicase ribozyme at hand (Attwater et al. 2010; Wochner et al. 2011).

There is another, comparably important criterion to be met for the evolution of the RNA-World towards the first living cell to proceed, and that one is *ecological* in nature. Assume that we have a few different self-replicating RNA replicase ribozymes and a constant supply of activated monomers from a geochemical source. The initial excess of resources (monomers) allows all the replicases to reproduce and establish their own populations. These exponentially increasing replicase populations will inevitably exhaust the constant monomer supply sooner or later, ultimately reducing the concentration of available monomers in the environment to a break-even level at which even the *fastest replicating* (i.e., fittest) replicase population stops growing. At the break-even resource level the populations of all other replicases are already *decreasing*, and they will continue doing so until they go extinct. This is the result of *competition* among the different replicase species for monomers, leading to the survival of the fittest, i.e., the victory of the most efficient replicase. According to the Gause principle of competition (Meszéna et al. 2006) the maximum number of coexistent replicator populations is equal to the number of different resources they exploit. If the replicators do not discriminate with respect to the different monomers that they use for their own replication, then there can be only a single winner. If the different monomer types (A, U, C, G for RNA) count as different resources, i.e., if different replicators use the monomers differentially, then the number of potentially coexistent replicators is equal to the number of monomer pairs (two in this case: A+U and C+G, (Szilágyi et al. 2013)). Since the differential use of monomers by the replicators – i.e., a marked difference in their A+U and C+G demand - would represent a severe constraint on their function (replicase activity), it seems reasonable to assume that the monomer pool constitutes a single resource, which implies a single winner. That is, for more than a single replicator species to coexist, some mechanism is needed that circumvents the problem of competitive exclusion, because further evolutionary improvements of the victorious replicase depend on its *cooperation* with RNA molecules helping its own population growth.

The straightforward ally could be a ribozyme catalysing a reaction that produces monomers from another geochemically supplied resource, i.e, a simple *metabolic enzyme*. The adoption of a single-step metabolism would be driven by the selective pressure on the replicase to exploit new resources present in the environment from which extra supplies of activated monomers can be produced. Ribozymes from the random-sequence replicator population of the RNA-World may be selected for the useful metabolic function and copied by the replicase, which in turn benefits from the increased monomer supply. This mutualistic interaction (cooperation) of the replicase and the metabolic ribozyme allows for a shift of the system towards *autotrophy* through the *construction of a new niche* that, thanks to their cooperation, becomes available for both the replicase and the metabolic ribozyme. The new niche is the potential to exploit the new compound – a resource thus far useless in replication – that the metabolic replicator is able to convert to activated monomers. The repeated inclusion of new metabolic ribozymes into the evolving RNA replicator community implies increasing autotrophy and metabolic efficiency of the reaction network through the process of *retroevolution of metabolism* ((Horowitz, 1945), Fig.2.). Our Metabolically Coupled Replicator System (MCRS) has been developed for studying the dynamical properties and the evolutionary

potential of such a community of cooperating ribozymes. The main questions to answer with the models are:

- Can a metabolically coupled set of replicators be coexistent in spite of the inevitable competitive interaction between the different replicator types? If so, under what environmental conditions does coexistence occur?
- How many metabolic replicators can be coexistent in MCRS?
- Can the MCRS resist the invasion of parasitic replicators which use the monomers and the service of the replicase for their reproduction but do not contribute to monomer production or replication at all?
- Can metabolically active replicators develop from a random replicator set?
- Is there any further evolutionary potential in MCRS through the acquisition and the development of new replicator functions?

2. General assumptions and the mean-field version of MCRS

Below we detail the basic assumptions of the MCRS model family, first specifying the mean-field version in which no spatial structure of the replicator community is considered. After showing that the mean-field model is not viable, we turn to the assumptions related to the spatial structure of the surface-bound RNA World, and specify the details of the spatially explicit core version of the MCRS scenario.

Assumption 1. *The chemical identity of early replicators*. The MCRS framework does not make explicit assumptions with respect to the chemical identity of prebiotic replicators, but straightforward general principles constrain the possibilities to modular (and, consequently, digital) structures capable of unlimited heredity (Szathmáry, 2006). These constraints practically exclude the majority of known chemical entities from among the plausible molecule types, except for variants of recent nucleic acids and proteins (Eschenmoser, 2007; Hall, 2004; Robertson and Joyce, 2010; Nielsen, 2009). Most researcher of the origin of life today agree that RNA, or RNA-like molecules are by far the most likely entities responsible for booting up life on Earth 3-4 billion years ago (Chen et al. 2007; Gilbert, 1986; Joyce, 2002; Robertson and Joyce, 2010). The MCRS is built on the RNA world scenario allowing for some chemical variations but maintaining the postulates of a modular, template-replicated macromolecule as the basic chemical entity of prebiotic evolution.

Assumption 2. *Error-free replication*. As explained earlier, one of the most difficult "missing links" in the MCRS scenario is that of RNA replication. The sequence of a relatively simple, yet sufficiently accurate RNA-dependent RNA polymerase ribozyme has not been discovered so far. Evolving such a replicase ribozyme is one of the biggest challenges for recent *in vitro* RNA evolution experiments (Attwater et al. 2010; Johnson et al. 2001; Rohatgi et al. 1996; Wochner et al. 2011). Lacking an efficient RNA replicase ribozyme we need to assume for the time being that the template replication of RNA molecules was nevertheless possible at the time of the wake of life. The most straightforward solution would be to suppose that there was a – so far undiscovered – replicase ribozyme present in the RNA world after all, which seems not to be unrealistic given the promising experimental results lately. A minor difficulty arises from the omission of the fact that any template replication is prone to mismatch errors (*mutations*) resulting in copies slightly different from the template. In fact this is the *error catastrophe* problem that the coexistence models of prebiotic evolution (i.e., the hypercycle, (Eigen and Schuster 1979), the stochastic corrector model (Szathmáry and Demeter, 1987), parabolic growth models (Szathmáry and Gladkih, 1989) and MCRS (Czárán and Szathmáry, 2000; Károlyi et al. 2002)) are meant to solve in the first place, but it is essentially circumvented by the assumption that the genetic information to be transmitted is split into short sequences. Therefore MCRS makes the simplifying assumption that RNA replication

is error-free on the ecological time scale for which the coexistence of metabolic replicators is investigated.

Assumption 3. *Double-stranded RNA*. Another difficulty related to the problem of experimental RNA replication is that even if the complementary strand can be formed, the copy cannot be separated from the template without imposing chemical conditions on the system that are very far from any reasonable assumption of prebiotic environmental conditions (Szathmáry, 2006; Patzke and von Kiedrowski, 2007). For lack of empirical knowledge on this issue we are again forced to assume that strand separation does occur somehow due to a mechanism so far unknown. As an initial simplifying assumption we assume that the sister strands of replicating RNA molecules are identical – an assumption that will be relaxed later (see Section 4).

Assumption 4. *Enzymatic activity of replicators*. Many different RNA molecules are known to take part in several vital biochemical processes of recent cells as catalysts (*ribozymes*, (Cech, 2009)). Early prebiotic RNA world systems must have relied mostly on the catalytic potential of ribozymes, because translation and thus more efficient protein enzymes are later achievements of evolution. The broad catalytic potential of RNA molecules was justified in different independent experimental studies (Bartel and Unrau, 1999; Chen et al. 2007; Landweber et al. 1998, Lilley, 2003).

Assumption 5. *Metabolism*. The key assumption of the MCRS is that each member of a set of different replicator types (i.e., replicator macromolecules of different nucleotide sequences) catalyses a single reaction in a hypothetical metabolic reaction network in which their own building blocks (*monomers*) are produced. Therefore monomers for replication are self-supplied only in the presence of a complete set of metabolic replicators (Fig.3.A.); any one of them missing halts monomer production altogether. Notice that we do not yet assume any explicit topology and stoichiometry for the metabolic reaction network here, even though it might be of substantial effect on the actual dynamics of the metabolic replicator system.

Based on these assumptions the mean-field version of the MCRS (Czárán and Szathmáry, 2000) model can be set up, in which the change of the frequencies (concentrations) of the metabolic replicators (f_i) are given as

$$\frac{df_i}{dt} = f_i(k_i \cdot M - \varphi(f)) \quad , \qquad \text{Eq. 1}$$

where k_i is the replicator-specific growth rate, $\varphi(f)$ is the outflow function which keeps the total concentration of replicators constant within the system, without altering their relative frequencies. $\boldsymbol{M}$ is the efficiency of metabolism, the network of chemical reactions in which each of the individual reactions is specifically catalysed by one of the metabolic replicators. Metabolic efficiency is calculated as the geometric mean of the replicator frequencies (concentrations) within the system:

$$M = \left(\prod_{i=1}^{n} f_i \right)^{\frac{1}{n}} \quad , \qquad \text{Eq. 2}$$

where n is the number of essential metabolic replicator types in the system. The metabolic function

316 ***M*** is the same for all the replicator types, because it represents the concentration of the product of
317 metabolism, i.e., the supply of monomers, which is the single common resource of self-reproduction
318 for all the replicators present in the system. Therefore the only parameter that determines the growth
319 rate of replicator *i* is its replication rate k_i in Eq.1. Consequence: the replicator of highest k_i
320 competitively excludes all the other ones and the metabolic community collapses in the mean-field
321 version of the MCRS model (Czárán and Szathmáry, 2000). In fact the system is exterminated
322 already by the exclusion of the *first* essential metabolic replicator type, because $f_i = 0$ for any *i*
323 implies $\boldsymbol{M} = 0$ in Eq.2. Note that Eigen and Schuster (Eigen and Schuster, 1979) had considered and
324 outright rejected a model of similar dynamics, precisely because it is not coexistent in a well-mixed
325 system.

326

327 ***3. The spatial version of the Metabolically Coupled Replicator System – the Metabolic Replicator***
328 ***Model (MRM)***

329 The disappointing conclusion of the mean-field model turns to its exact opposite with the
330 assumption that the MCRS is bound to a mineral surface, so that the interactions of the replicators
331 (metabolic cooperation and competition for monomers) become locally context dependent.

332 Experimental data of very different sorts provide strong indirect support for the idea:
333 mineral underwater surfaces (rocks of pyrite, clay minerals like montmorillonite, etc.) can be
334 catalysts for nucleotide binding (Ferris, 2006, Ferris et al. 1999); they might be responsible for the
335 homochirality of biomolecules (Hazen et al. 2001, Joshi et al. 2011); they are supposed to have
336 assisted membrane production and thus the formation of the first proto-cells (Hanczyc et al. 2007);
337 and they may have protected replicators from the harmful effects of UV radiation (Biondi et al.
338 2007).

339

340 ***3.1. Space-related assumptions of the spatially explicit MCRS model***

341 **Assumption 6.** *Replicators are bound to mineral surfaces.* The most probable arena for prebiotic
342 replicator evolution may have been on mineral surfaces which can bind RNA molecules reversibly
343 through divalent cations (Franchi et al. 2003). Detachment and re-attachment of parts of the
344 macromolecules result in their caterpillar-like movement on the surface, which is in turn responsible
345 for their limited rate of spatial mixing – a feature that later will be shown to be of crucial
346 importance for their population dynamics. The two-dimensional arena is a lattice of binding sites,
347 each site harbouring a single replicator at a time. Replicator movement is represented by swapping
348 the contents of neighbouring sites. We specify the details of replicator movement in Section 5.

349 **Assumption 7.** *Initial replicator diversity generated by spontaneous polymerisation.* Surface-
350 catalysed RNA polymerisation results in a diverse pool of oligo- and polynucleotides of different
351 lengths and random nucleotide sequences (Copley et al. 2007; Garay, 2011, Ma et al. 2007;
352 Manrubia and Briones, 2007). This random community of replicators is then selected for useful
353 metabolic functions contributing to monomer production.

354 **Assumption 8.** *Local metabolic interactions on the surface*. The limited mobility of replicators on
355 the mineral surface makes their metabolic and competitive interactions local. Local metabolic
356 interactions mean that the metabolite molecule produced by a ribozyme replicator needs to be
357 delivered to the ribozyme catalysing the next reaction of metabolism before the metabolite decays
358 or desorbs from the surface. This requires that the corresponding metabolic replicators be
359 sufficiently close to each other in space. As an implicit proxy to this criterion we assume that all the
360 metabolically essential replicators need to be present within a certain area called the *metabolic*
361 *neighbourhood* (Fig.3.C.) around a replicator so that it has a sufficient local monomer supply for its
362 replication. This corresponds to the local application of Eq. 2 within each metabolic neighbourhood
363 instead of the whole replicator community.

Assumption 9. *Surface diffusion of metabolites*. The detailed chemical nature of precursors, intermediary metabolites and monomers is disregarded in the MCRS, just like the topology of the metabolic reaction network itself. What we implicitly consider are a few general features of small molecules in relation to their movement on and detachment from the mineral surface. We assume that small molecules move on the surface faster than macromolecules do, and they can desorb from the surface with a probability higher than replicators. Both of these assumptions reflect that small molecules (e.g. monomers) are certainly less attached to the surface than macromolecules.

Assumption 10. *Local competition for monomers*. Like metabolic interactions, competition is also local in the spatially explicit MCRS model: replicators within the *replication neighbourhood* (cf. Fig.3.C.) of an empty site compete for the possibility to put a copy of themselves onto the focal empty site. The chance of replicator I_i to win depends on its replication parameter k_i and its local monomer supply M_i.

3.2. The stochastic cellular automaton implementations of MCRS

Basic model setup. The computer implementation of the Metabolically Coupled Replicator System scenario is a series of stochastic cellular automaton (SCA) models: the Metabolic Replicator Model (MRM) family. A set of n different, metabolically active ribozyme replicators are assumed to compete for the monomers which they produce themselves in cooperation, through catalysing the reactions of a simple metabolism (Fig.3.A.). Each replicator occupies a site of the SCA lattice representing the mineral surface on which all the interactions take place. The opposite margins of the lattice are merged forming a toroidal structure to avoid edge effects. The number of possible states for a site is $n + 1$, including the "empty" state and the n different occupied states. The lattice size we used throughout the simulations was 300 x 300, which is sufficiently large to avoid strong periodic effects but is still manageable in terms of computer resources. One generation (from t to $t + 1$) consists of elementary updates equal in number with the number of sites in the lattice (90.000). The updating algorithm is random: the state of each site is updated once per time unit on average, in a random order *(asynchronous updating rule).*

Update processes: replication and decay. Empty and occupied sites are updated by separate algorithms. Occupied sites turn to the "empty" state (*replicator decay*) with the constant replicator decay probability p_d. "Empty" sites can become occupied by a copy of one of the replicators from within the *replication neighbourhood*; the replicators there compete for the focal empty site. The chance of a replicator to win the competition and put a copy of itself to the empty site depends on its replication parameter and the local monomer supply within the *metabolic neighbourhood* of the focal replicator (Fig.1.C.). The size of the metabolic neighbourhood is considered proportional to the average distance that a small molecule (metabolite or monomer) can cover by surface diffusion before it either desorbs from the surface or is consumed in a replication process. The individual "claim" C_f of the replicator f for occupying the empty site depends on its monomer supply M_f and its specific replication rate k_f as

$$C_f = k_f \cdot M_f \quad , \qquad \text{(Eq. 3)}$$

and

$$M_f = \sqrt[n]{\prod_{i=1}^{n} x_i(f)} \quad , \qquad \text{(Eq. 4)}$$

where $x_i(f)$ is the number of type replicator i within the metabolic neighbourhood of the focal

replicator *f*, and *i* runs through all replicator types needed to catalyse the metabolic reactions (i = 1, ..., n). Thus, the local monomer supply of the focal replicator *f* depends on the presence of *all* metabolic replicators within its own metabolic neighbourhood – with any one of the n metabolic replicator types missing the corresponding $x_i(f) = 0$ and thus also $M_f = 0$. This in turn implies no local monomer production and therefore no chance of replication for the focal replicator *f*. Each replicator within the replication neighbourhood of an empty site has a chance to occupy the empty site with a copy of itself:

$$p_f = \frac{C_f}{C_e + \sum_m C_m} \quad , \qquad \text{(Eq. 5)}$$

where m runs through all replicators within the replication neighbourhood of the focal replicator *f*, and C_e is a constant representing the claim of the empty site for remaining empty. Obviously, the probability that the empty site remains empty is

$$p_e = \frac{C_e}{C_e + \sum_m C_m} \quad . \qquad \text{(Eq. 6)}$$

Note that in the basic MRM there is no specialised replicase replicator in the system. It is implicitly supposed here that the replicase "service" is supplied either by the mineral surface itself, or by a very rudimentary replicase ribozyme which is present in excess on the surface. This implicit assumption will be relaxed later by the explicit inclusion of a replicase ribozyme.

Replicator diffusion. The movement of replicators on the mineral surface is implemented using the Toffoli-Margolus algorithm: randomly chosen 2x2 blocks of sites are rotated by 90° left or right with equal (0.5) probability (Toffoli and Margolus, 1987). The intensity of replicator diffusion is scaled by the average number D of diffusion steps per site per generation. Note that even $D = 0$ represents some minimum mixing of replicators on the surface, due to the fact that each newborn copy is placed into a site different from – adjacent to – the one occupied by the parent (template).

Structured (porous) habitat. The basic model was also modified to account for the dynamical effects of an *ab ovo* compartmentalised, i.e., structured, habitat. Many of the possible minerals on which prebiotic replicator evolution might have taken place are in fact of a porous structure (Fig.4.). The pores, which are connected by capillary channels, represent compartments relatively separated from other pores. In this spatially structured version of the MRM the pores take the role of interaction (metabolic and replication) neighbourhoods: each pore is considered as an open stirred-tank reactor connected by the in- and outflow of small molecules and, occasionally, of macromolecular replicators as well. Each pore can support a certain number of replicators (*pore capacity*), and the concentrations of small molecules ("resources" and "monomers") are explicitly followed. A given fraction of streaming small molecules can dock within the pore reducing pore capacity. The metabolic efficiency (M) of a pore is calculated based on its replicator and monomer contents, taking the monomer threshold for replication into account. Since the structured model is more explicit in terms of chemistry, and somewhat different in terms of spatial structure compared to the basic model, it is of very high importance to evaluate its predictions against those of the basic MRM.

4. Spatially explicit simulations

The most striking result of the spatially explicit models of the MCRS scenario is that all implementations are very robustly coexistent within a broad range of their space-related parameters. This suggests that local interactions among a set of metabolically essential ribozyme replicators are sufficient to maintain their cooperation and to neutralise, or at least to reduce, the competitive effects which drive the mean-field system (cf. Section 2) to extinction. A typical run of the non-structured simulation yielded the time series on Fig. 5, with 4 metabolic replicators of substantially different replication parameters k_i.

4. 1. The ecology of the spatial models

The space-related parameters of the basic MRM which are relevant for the coexistence of the metabolic replicator community are: 1) metabolic neighbourhood size, 2) replication neighbourhood size and 3) the diffusion parameter of the replicators. Fig.6. summarizes the results of a series of simulations scanning through the space of these three parameters.

What explains the fundamental difference in the dynamics of the mean-field model and the spatial models? The answer is that in the spatial model the local range of metabolic cooperation gives an indirect advantage to *rare* replicator types through local metabolism, because their metabolic neighbourhoods are easily complemented by the more common types, therefore they have a better chance for replication than the common types, most of which lack the presence of the rare type within their metabolic neighbourhood. The larger the difference in frequency between two replicator types the larger the advantage of rarity.

Eq.3. implies that the fitness of a replicator (I_f) consists of two components – a *direct* and an *indirect* one. k_f, the specific and constant replication parameter is the direct fitness component: low k values provide few opportunities for replication, thus the density of the corresponding replicator type in the community is low (the replicator is rare). The advantage of the rare type comes from the indirect fitness component (M_f), and it acts through the better local monomer supply of the rare types on average. The complete metabolic replicator community is coexistent when the fitnesses of all the replicator types are equal: $C_i = C_j$ for any (i, j). That is, replicators of low direct fitness compensate for their handicap by a higher indirect fitness. Since the indirect fitness component increases with rarity, the negative feedback of replicator population density on fitness regulates the community to coexistence in a broad range of the parameter space.

Obviously, the spatial parameters of the model (metabolic neighbourhood size, replication neighbourhood size and replicator mobility) affect coexistence through their effects on the indirect components of replicator fitnesses. Let us consider these in turn.

Metabolic neighbourhood: The size of the metabolic neighbourhood is a proxy to the distance that metabolites and monomers travel by surface diffusion before disappearing either by desorption or by reaction (cf. Assumption 8). Very small metabolic neighbourhoods mean a very localised metabolism, which translates to a strong advantage of rarity: low frequency metabolic replicators with very small metabolic neighbourhoods have better chances for replication than common ones, because their indirect fitness component is very high. Increasing the metabolic neighbourhood shifts the system towards the mean-field approximation; in the limit case of the metabolic neighbourhood being equal to lattice size (300x300) we arrive at the mean-field model which we know to go extinct (cf. Section 2).

Replication neighbourhood: The replication neighbourhood of an empty site corresponds to the distance to which the „offspring" of a replicator can be placed from its parent. Common sense suggests that this should not be large, because it is difficult to imagine the mechanism which could put the copy far from the template in spite of the relatively strong adherence of both to the surface. A long-distance movement by the copy would require its detachment from, and then its distant re-attachment to the surface – a very unlikely series of events indeed. Even so, increasing the size of the replication neighbourhood has an obvious mixing effect: it decreases the probability that the

offspring remains close to the parent, i.e., the chance of aggregated pattern formation decreases. Note, however, that increasing the replication neighbourhood *does not* shift the system towards the mean-field case, because the metabolic advantage of rarity remains the same.

Replicator mobility (diffusion): Faster replicator movement means better mixing, too. It is obviously advantageous for the coexistence of the metabolic replicator community, especially if the metabolic neighbourhood is small. Less mixing would lead to the aggregation of conspecific replicators, which would drastically decrease the chance of metabolic complementation on the spatial scale of local metabolism (i.e., at the scale set by metabolic neighbourhood size). Replicator mobility (diffusive mixing) increases the overall fitness of the community by increasing the number of complete metabolic neighbourhoods, and thus the indirect fitness of *all* replicators in the system.

The combined effects of these three spatial parameters on the stationary states of the MRM system are shown on Fig.6. (Könnyű and Czárán, 2013). The best conditions for the coexistence of metabolically coupled replicator communities are at relatively small metabolic neighbourhood sizes and intensive replicator mixing (the latter condition seen at large replication neighbourhoods and/or high replicator mobility).

The parameters of the spatially structured „pore-model" analogous to metabolic neighbourhood size and replicator mobility in MRM are pore size (i.e., the maximum number of replicators fitting into a pore) and replicator migration, respectively. Fig.7. shows that within the coexistent section of the space of these two parameters the trend in the pore-model is the same as in MRM: larger pore size decreases, whereas more replicator mobility increases the fitness (and the mean density of the replicator community). Since the pore model is more explicit in terms of chemical detail (i.e., it considers the constant input of a „resource compound" which can be converted to monomers by the replicator community of a pore, provided it is metabolically complete), the convergence of the results of the two models is encouraging.

The original problem which MRM (and the hypercycle model) intended to solve is the maintenance of genetic information surpassing the error threshold and sufficient to code for a machinery complicated enough to be capable of its own reproduction (Eigen and Schuster, 1979; Kun et al. 2005; Maynard-Smith, 1979; Niesert, 1987; Niesert et al. 1981; Takeuchi and Hogeweg, 2007). Considering this problem as the central one, there is another parameter of the MRM of crucial importance: the number of different replicator types that the model can keep coexistent, i.e., the maximum attainable system (genome) size (n). One simple constraint is trivial: system size cannot exceed the maximum number of replicators fitting into the metabolic neighbourhood, or else a complete local metabolism is impossible, so larger metabolic neighbourhoods should be able to harbour larger systems. However, increasing metabolic neighbourhood size decreases the advantage of rarity at the same time; therefore we expect the largest possible viable systems to be maintained at intermediate metabolic neighbourhood sizes. The actual attainable system size is also limited by the level of spatial mixing – the more intensive it is, the larger the biggest sustainable system should be. We have tested maximum viable system size as the function of space-related model parameters both in MRM and in the pore-model (Fig. 8.), and the results confirm these expectations: high mixing and intermediate metabolic neighbourhood sizes allow for the coexistence of over 10 replicators. Towards the limit of infinite diffusion the Metabolic Replicator Model approaches Wilson's trait group mechanism of coexistence (Maynard-Smith and Szathmáry, 1995; Szathmáry, 1992; Wilson, 1975).

4.2. Parasites, complementary strands and facultative cooperators in MRM

The metabolic replicator system, like any cooperative community, is exposed to "cheaters", i.e., individuals taking advantage of cooperation by others, but not investing into cooperation themselves. Such free-riders enjoy the fitness advantage of reduced resource investment compared

to cooperators, and they spread in the community until cooperation breaks down altogether. This is what happens in parasite-infected cooperative communities without a proper reward/punishment scheme in effect. Intentional rewarding or punishment is, of course, out of question in macromolecular communities. The cooperating members of a hypercyclically coupled replicator community have no means of feeding back the damage from parasitism to the parasite itself. This is why the naked hypercycle is doomed to collapse upon the emergence of mutant replicators acting as selfish or shortcut-parasites (cf. Introduction).

The only conceivable parasite of the MCRS is one that uses up monomers for its replication but does not contribute to monomer production (Fig.3.B.). Any mutant failing to contribute to the common good is a parasite; therefore we expect a whole range of different parasites – a parasitic quasispecies - to emerge in any metabolic replicator system. Neglecting the slight differences in their dynamically relevant parameters we lump the members of the parasitic quasispecies into a single replicator category. The metabolic cooperation mechanism of the surface-bound MCRS provides an "automatic" delivery of efficient punishments to such selfish parasites: wherever they pop up and start spreading, monomer production is impaired, which in turn locally stops the replication of all replicators including the parasite. Since extinctions occur only where parasites prevail, local extinction decimates the parasite more than cooperators. This mechanism is sufficiently powerful to keep the parasitic quasispecies in check even if it has the highest replication parameter in the community (Fig.9.A.). High parasite replicability is a feasible assumption, since no enzymatic function constrains the secondary structure of a parasite: it will be selected for fast reproduction, i.e, it should be short and loosely folded.

The most relevant parameter of the MRM with respect to its parasite resistance is not the replication rate of the parasite, but replicator mobility. At limited replicator mobility even a very fast reproducing (of large k_p) parasite will attain a low steady state frequency in the MCRS, but at high mobilities the parasite can destroy cooperation (Branciamore et al. 2009; Czárán and Szathmáry, 2000). Since very high mobilities are not reasonable to assume for surface-bound macromolecules, this extreme case does not constrain the feasibility of the model. The feasible space-related parameter range of a viable MRM exposed to parasitic mutants is therefore small to moderate metabolic neighbourhood sizes and small to moderate replicator mobilities. We did not explicitly tackle the dependence of diffusibility and desorption rate on replicator length in this model, but we may safely assume that short parasitic replicators move faster and desorb easier from the surface than longer, metabolically active replicators do. These two effects of sequence shortening are assumed to quench each other: one is advantageous and the other is detrimental for parasite persistence.

Note that the parasitic quasispecies can be rather heterogeneous with respect to length and replication parameter, but for the cooperating replicator community the only dynamically relevant feature of a parasite is its being a cheater (i.e., that it does not contribute to monomer production). The replicability of the parasite is quite irrelevant for the cooperators, as they can repress them through the metabolic "punishment" mechanism anyway. The difference in the replication rates of different parasites plays a role only in inter-parasite competition: the fastest replicating type of the parasitic quasispecies excludes all the other types (Könnyű and Czárán, 2013), in perfect accordance with the Gause principle ((Meszéna et al. 2006), Fig.9.B.). Thus the outcome of a typical MRM + parasitic quasispecies simulation is the coexistence of all cooperators and the fastest parasite, with the latter attaining a low and steady equilibrium frequency in the community.

Two special modifications of the MCRS model deserve mention here, because their dynamical consequences are somewhat similar to that of introducing parasites. The first such modification is relaxing the template/copy identity postulate (Assumption 3). RNA template and copy strands are not identical but complementary in their nucleotide sequences, and possibly very

different in secondary structure (except for palindromes, (Boza et al. 2014; Ivica et al. 2013)). The complementary strand (i.e., the "gene") of a metabolically active ribozyme is, in all probability, functionally inactive, which makes the copy of the ribozyme similar to a parasite from an ecological point of view: it consumes monomers, but it does not contribute to producing them. The difference is that the copy is the offspring of a ribozyme, and the copy of a copy is a functional ribozyme again, which is not the case with a real parasite. Assuming that the functional (ribozyme) forms have lower replicability than the complementary strands (Ivica et al. 2013; Könnyű and Czárán, 2014) because of their – presumably more compact – secondary structure, the system behaves like the MRM + parasite quasispecies model, except that the "gene" copies do not exclude each other: all the complementary strands coexist with the metabolically active ribozyme forms. The MCRS proved to be viable in this pheno/geno version as well (Fig.10.A.), which is not a big surprise given the robust parasite resistance of the original MRM (Könnyű and Czárán, 2014).

The other special modification of the model is the inclusion of *facultative metabolic cooperators*: replicators which increase the efficiency of metabolism, but are not essential for monomer production. The metabolic benefit provided by a facultative cooperator may come, for example, from its acting as a co-factor of another, essential metabolic ribozyme. Such facultative cooperators are very similar to parasites in their dynamical properties, except that their negative effect of diluting the local assembly of essential ribozymes is counteracted by their positive effect on metabolic efficiency. Of course they also coexist with the original MRM, and exclude other parasites which do not help metabolism (Könnyű and Czárán unpub.) (Fig.10.B.).

5. Adaptive evolution in MRM

The stable coexistence of the core of MRM (the metabolically essential replicator set) with a non-functional parasitic replicator is the most important feature of the surface-bound MCRS from the viewpoint of its evolvability. The benefit of the presence of a parasitic replicator lies in its pre-adaptive value: it remains persistent in the functioning MCRS without causing much damage, and it can freely mutate to obtain new functions potentially increasing the fitness of the community of cooperating replicators. The most straightforward adaptive enhancements of the system may advance through improvements of the existing metabolic ribozymes, simply by selection towards better, or more specialised, enzymatic activities.

5.1 Adaptations improving metabolic efficiency

Better catalyst may drive better metabolism, and local replicator communities fed by more efficient metabolism will obviously displace others from the surface by competition. However, adaptations towards better metabolic functions are necessarily traded off with replicability: more efficient ribozyme structures tend to be more compact, therefore they are also more difficult to unfold and replicate. On the other hand, less efficient enzymes can be more versatile in terms of possibly catalysing more than a single reaction of metabolism, if two (or more) different, but energetically similar foldings of the macromolecule are possible, and each has some catalytic activity with respect to a metabolic reaction different from those of the other foldings. Such "promiscuous" catalytic activities have been reported both for protein (Khersonsky and Tawfik, 2010; O'Brien and Herschlag, 1999) and for RNA (Ancel and Fontana, 2000;Schultes and Bartel, 2000) enzymes. The different catalytic effects of promiscuous ribozymes are also in a trade-off relation one with the other, for at least two different reasons. First, spending time in one of the foldings means that the other folding is inactive, i.e., the ribozyme works in time-sharing mode. Second, the fact that the molecule is able to trans-fold to other secondary structures implies that none of its secondary structures is very stable (compact): a handicap with respect to its catalytic activity in each folding. These trade-off constraints are reflected in the following supplementary assumption applied in the next modification of MRM:

Assumption 11. *Trade-offs in replicator features*. We assume two-way trade-offs among the features of metabolic ribozymes with potentially two different catalytic activities. The first trade-off is between the two enzymatic activities, the second one is between the enzymatic activities and the replicability of the metabolic replicator molecule. These trade-off constraints restrict the available combinations of the three features below the trade-off surface shown on Fig.11 (Könnyű and Czárán, 2011). The parameters *b* and *g* of the trade-off function scale the strength of the trade-off between the two enzymatic activities and the enzymatic activities and replicability, respectively.

Replacing the numbers of type *i* ribozymes $x_i(f)$ with the *total type i activity* of the different replicators *k* ($\sum_{k=1} E_{i,k}(f)$) within the metabolic neighbourhood of the focal replicator *f* leads to the substitution of Eq.4 with

$$M_f = n\sqrt{\prod_{i=1}^{n}\sum_{k=1} E_{i,k}(f)} \quad , \qquad \text{Eq. 7}$$

in which each metabolic replicator *k* within the given metabolic neighbourhood counts with *only one* of its enzymatic activities *i*, which is drawn at random with weights of choice proportional to the actual activities $E_{i,k}$ of replicator *k*. If the focal replicator *f* is the one copied from among the candidates in the replication neighbourhood, then the copy is either identical with its template or – with a small probability – it is a mutant with its catalytic activities and replicability constrained by the trade-off surface, but otherwise chosen at random.

Depending on the shape of the trade-off function of catalytic activities (parameter *b*), and on replicator mobility *D*, the simulations reveal two different outcomes: the system ends up either in a dominantly "specialist" replicator community consisting of single-activity metabolic ribozymes, or in the dominance of "generalists", i.e., catalytically less efficient but bifunctional replicators (Fig.12.). The criteria for ribozyme specialization proved to be hard trade-off (low values of *b*) between the two catalytic activities, and moderate replicator mobility – both criteria falling in the most feasible zone of the parameter space. Hard trade-off means that the sum of the two enzyme activities of "generalists" (bifunctional ribozymes) is less than that of any of the two "specialists". Of course more mixing (larger *D*) is beneficial for specialization, because it prevents the aggregation of identical templates and copies, which prevents metabolic complementation of specialists. Note that parasites (replicators with both of their catalytic activities next to zero, but with very high replicability) are kept at very low frequencies in this model implementation, just like in all previous ones, provided that replicator mobility is not extremely high (Könnyű and Czárán, 2011).

5.2. The evolutionary acquisition of new functions by the metabolic replicator community

Besides improving the catalytic activities of existing members of the metabolic replicator community as explained above, even more innovative adaptations might come from adopting new functions by mutants of either a core replicator or the parasite of the system. The mutant may be a new metabolic ribozyme possibly opening a new, more efficient chemical route to monomer production, but it may obtain other functions increasing the fitness of the cooperating replicator community in radically different ways. Such adaptations may, for example, accelerate the replication process itself through improving the replicase ribozyme, thus increasing the fitness of each replicator in the community; or they might contribute to the completion of the system with the third essential infrabiological (Szathmáry et al, 2005) component of life: the membrane envelope.

Replicase evolution: For evolutionary adaptation to take place within the MCRS, template replication has to work one way or another. This requires the replicase function to be available from the outset, either as a service of the environment (some mineral surfaces are shown to have a basic catalytic activity helping spontaneous RNA template replication – (Ferris, 2006)) or in the form of a

simple replicase ribozyme of the initial random replicator population capable of copying itself and other RNA replicators (Fig.13.A. (Könnyű et al. 2008)). Assuming that the metabolic replicator community has already domesticated a parasite for the replicase function, that replicase can mutate to become worse or better in its role. Allowing for mutations both in the negative and the positive direction of replicase activity we studied the evolutionary dynamics of the core MRM + replicase system, based on the following supplementary assumption:

Assumption 12. *Beneficial and deleterious mutations of the replicase*. Parasites can mutate to obtain increasing template replicase activity (beneficial for MCRS) or replicase inhibitory effects (deleterious for MCRS). Both these types of mutation are traded off with the replicability of the parasite: the stronger the functional effect of the mutation (in any direction), the smaller the direct fitness component (k) of the replicator.

Simulations reveal that inhibitory parasites disappear from the system, because 1) they kill off metabolic replicators locally more efficiently, thus committing suicide faster, or 2) they evolve to higher replicability to the expense of their inhibitory effect. On the other hand, beneficial mutants spread and achieve substantial frequencies (Fig.13.B.) in the replicator community, provided that the trade-off relation of replicase activity and replicability is not very rigid, and replicator mobility is not too high. Moreover, the overall density of the replicator community also increases as higher replicase activity builds up, indicating the evolutionary benefit of improving an aspecific replicase function to the whole system.

6. Perspectives of the MCRS approach

First a few comments on the relationship of the systems surveyed here to Gánti's various suggestions of chemical supersystems are in order. The 'classis' 1971 Hungarian edition of the Principle of Life (Gánti, 1971) coined the term 'chemoton', but then it referred to the system doublet made of an autocatalytic metabolic cycle and a template replicator only. In this sense we have also dealt with similarly organized infrabiological systems (Szathmáry et al, 2005). But there is a crucial difference: the idea of hereditary catalytic effects by the templates entered Gánti's thinking only towards the end of the seventies only (Gánti, 1978), and then strictly within the fully-fledged metabolism-boundary-genetic material tripartite systems that is nowadays being referred to as the chemoton We think, however, that catalytic reaction channelling must have been an indispensable feature of any chemical supersystem maintaining even a minimal metabolism and capable of self-reproduction (Deamer and Weber, 2010; Meléndez-Hevia et al. 2008; Orgel, 2000, 2004; Pross, 2004; Szathmáry et al, 2005). Therefore, the most important agents of an early metabolism-replicatorsystem must have been the catalysts which, through the metabolism they drive, can produce their own building blocks, using externally supplied raw materials. The early RNA-World hypothesis provides an excellent starting point for a feasible scenario of the origin of life, because RNA is involved in two of the three infrabiological functions of life: a wide spectrum of catalytic activities for driving practically any *metabolism*, and a large variability of template-complementary module (nucleotide) sequences to attain *unlimited heredity* and self-reproduction (Chen et al. 2007;Gilbert, 1986; Joyce, 2002; Landweber et al. 1998).

Evolution requires reproduction, i.e., self-copying of sufficient accuracy. Even though we cannot yet pinpoint the agent which could have been able to copy itself at the earliest stages of chemical evolution, RNA, or RNA-like macromolecules are the primary suspects for this role as well. Recent laboratory experiments are very promising, with their results getting ever closer to the discovery of a self-replicating ribozyme, i.e., an RNA replicase that copies diverse RNA molecules of at least its own size with a sufficiently low mutation rate. Once we have that, the stage is ready for the evolution of an increasingly complex metabolism supplying monomers for the replicase

744 population, through the sequential adoption of other replicators playing the roles of metabolic
745 enzymes in a metabolic reaction network that becomes increasingly autotrophic (*retroevolution*,
746 (Szathmáry, 2007)). This process is driven by the ecological pressure towards occupying (or
747 constructing) new niches: the inclusion of new compounds supplied by the environment for
748 metabolism as old external resources become exhausted, one after the other, by the exponentially
749 increasing replicator population.
750 The other ecological constraint on the dynamics of the evolving replicator community is the
751 avoidance of competitive exclusion of any of the ribozymes playing a vital role in maintaining
752 metabolism and replication. The metabolically coupled replicator system (MCRS) model was
753 developed to demonstrate that this is possible, if the system is bound to a mineral surface, thereby
754 increasing the viscosity (i.e., limiting the spatial mixing) of the interacting replicators. The MCRS
755 model resists parasitic replicators in the sense that, under physico-chemically reasonable
756 assumptions, parasites cannot kill the system, even though they remain persistent at low
757 frequencies. Deleterious mutants of either the metabolic cooperators or the parasites are doomed to
758 extinction, because by hindering obligatory cooperation they decimate neighbouring cooperators
759 and thus cut their own monomer supply – in effect, they behave as suicide bombers, and go extinct.
760 The MCRS model framework may be improved in two main directions: in depth, by
761 explicitly considering important chemical details that have not been addressed so far; and in
762 extension, by broadening the approach to include completely new directions of MCRS evolution.
763 Some in-depth variants of MRM are being studied already: besides the "pheno-geno" version
764 considering replication to produce complementary strands ((Könnyű and Czárán, 2013), cf. Section
765 4.2), simple explicit metabolic reaction topologies with explicit metabolite and monomer
766 production and diffusion have been shown to work (Kőrössy et al, *unpub*.). These simulation studies
767 also show the limits of the surface-bound MRM. The size and the topology of the metabolic
768 network, just like the number of possibly coexistent replicators, are constrained mainly by the same
769 spatial factors that make the system work: the local nature of interactions, and the limited range of
770 the surface diffusion of replicators and metabolites. The limits set by these spatial constraints can be
771 pushed further out only by extending the MRM approach to new mechanisms of selection. The key
772 to such modifications is the unavoidable, but rarely fatal, presence of parasitic replicators in the
773 cooperating replicator community.
774 Neutral mutants of persistent parasites are free to random-walk across the sequence space
775 and may find functions beneficial for the cooperating replicator community. Such converted
776 parasites can be adopted by the system and might radically increase its fitness, by opening new,
777 efficient metabolic routes, improving replication (cf. Section 5.2), or producing membranogenic
778 molecules and trans-membrane channels.
779 Membrane production is the critical step towards the occurrence of the first protocell,
780 allowing for a new organizational level to occur, and new mechanisms for its evolution. Acquiring
781 the ability of membrane synthesis could provide the replicator community with *individuality*, of
782 profound evolutionary consequences. Autonomous membrane production could be achieved
783 through some mutant parasites evolving to ribozymes catalysing the production of membranogenic
784 (amphipathic) molecules from other metabolites, and the spontaneous insertion of their product into
785 the expanding membrane (Fig.14.). Encapsulating a replicase-aided MCRS into self-supplied
786 membrane compartments would establish a more effective, new level of selection for further
787 evolution of the system – it would be the organizational level of the *protocell*. The stoichiometric
788 coupling of membrane production to metabolism ensures the synchrony of doubling metabolite
789 content and membrane surface, which warrants the possibility of protocell fissions maintaining the
790 original volume/surface ratio through indefinitely many generations. Once in place, the membrane
791 capsule can adopt selective permeability functions or even active pumping of resource compounds
792 into the protocell, by evolving specific membrane-bound ribozymes (Khvorova et al. 1999).
793 Through such adaptations the protocell could achieve independence from the mineral substrate and

enter a new evolutionary regime: that of the internal reorganization of genetic, metabolic and transport functions, towards the cellular state as we know it in recent organisms. The simultaneous (co-)evolution of the genetic, the metabolic and the membrane subsystems could have occurred through the progressive sequestration scenario (Szathmáry, 2007), with metabolism becoming more complex, membrane channels more selective and genetic material organized in chromosomes. Modelling the early phases of protocell evolution along these lines is the intended direction of our future extensions to MRM; Fig.14. is a caricature of the idea, the model implementation of which is a task for the future.

Competing interests

The author(s) declare that they have no competing interests

Authors' contributions

TC, BK and ESz designed, analysed and interpreted the introduced studies. All authors contributed to writing the manuscript and approved the final version.

Acknowledgements

TC and BK acknowledge financial support from the Hungarian Research Foundation (OTKA Grant No. K100806). E. Sz. was supported by European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement No. [294332] and partly by EU COST action CM1304 "Emergence and Evolution of Complex Chemical Systems."

References

Ancel, W.L., Fontana, W., 2000. Plasticity, evolvability, and modularity in RNA. J. Exp. Zool. (Mol. Dev. Evol.) 288, 242–283.

Attwater, J., Wochner, A., Pinheiro, V.B., Coulson, A., Holliger, P., 2010. Ice as a protocellular medium for RNA replication. Nature Commun. 1, 76.

Bartel, D.P., Unrau, P.S., 1999. Constructing an RNA world. Trends. Genet. 15, M9–M13.

Biondi, E., Branciamore, S., Maurel, M.C., Gallori, E., 2007. Montmorillonite protection of an UV-irradiated harpine robozyme: evolution of the RNA world in a mineral environment. BMC Evol. Biol. 7, Suppl 2.

Boerlijst, C., Hogeweg, P., 1991, Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. Phisyca D. 48, 17–28.

Boerlijst, C., 2000. Spirals and spots: novel evolutionary phenomena throuogh spatial self-structuring. In Dieckmann U., Law R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions, IIASA and Cambridge University Press, Cambridge pp. 171–182.

Boza, G., Szilágyi, A., Kun, A., Santos, M., Szathmátry, E., 2014. Evolution of the division of labor between genes and enzymes. PLoS Comput. Biol. 10, e1003936, doi:10.1371/journal.pcbi.1003936.

Branciamore, S., Gallori, E., Szathmáry, E., Czárán, T. 2009. The origin of life: chemical evolution of metabolic sytem in a mineral honeycomb? J. Mol. Biol. 69, 458-469.

Bresch, C., Niesert, U., Harnasch, D. 1980. Hypercycles, parasites and packages. J. Theor. Biol. 85, 399–405.

O'Brien, P.J., Herschlag, D., 1999. Catalytic promiscuity and the evolution of new enzymatic activities. Chem. Biol. 6, R91–R105.

Cech, T.R., 2009. Crawling out of the RNA World. Cell. 136, 599–602.

Chen, I.A., Walde, P., 2010. From self-assembled vesicles to protocells. Cold Spring Harb. Perspect. Biol. 2, a002170.

Chen, X., Li, N., Ellington, A.D., 2007. Ribozyme catalysis of metabolism in the RNA world. Chem. Biodiv. 4, 633–655.

Cleaves, H.J., Ii, H., Scott, M.A., Hill, F.C., Leszczynski, J., Sahai, N., Hazen, R., 2012. Mineral-organic interfacial processes: potential roles in the origins of life. Chem. Soc. Rev. 41, 5502–5525.

Copley, S.D., Smith, E., Morowitz, H.J., 2007. The origin of the RNA world: Co-evolution of genes and metabolism. Bioorg. Chem. 35, 430–443.

Czárán, T., Szathmáry, E., 2000. Coexistence of replicators in prebiotic evolution. In Dieckmann U., Law R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions, IIASA and Cambridge University Press, Cambridge pp. 116–134.

Deamer, D.W., Weber A.L., 2010. Bioenergetics and life's origins. Cold Spring Harb. Perspect. Biol. 2, a004929.

Ehrenfreund, P., Cami, J., 2010. Cosmic carbon chemistry: from the interstellar medium to the early Earth. Cold Spring Harb. Perspect. Biol. 2, a002097.

Eigen, M., Schuster, P. 1979. The hypercycle, first ed. Springer-Verlag, Berlin.

Eschenmoser, A., 2007. The search for the chemistry of life's origin. Tetrahedron. 63, 12821-12844.

Ferris, J.P., 2006. Montmorillonite-catalysed formation of RNA oligomers: the possible role of catalysis in the origins of life. Philos. T. Roy. Soc. B. 36, 1777–1786.

Ferris, J.P, Hill, R.A., Liu, R., Orgel, L.E., 1996. Synthesis of long prebiotic oligomers on mineral surfaces. Nature. 381, 59–61.

Franchi, M., Ferris, J.P., Gallori, E., 2003. Cations as mediators of the adsorption of nuclec acids on clay surfaces in prebiotic environments. Origins Life Evol. B. 33, 1–16.

Garay, J., 2011. Active centrum hypothesis: the origin of chiral homogenity and the RNA-World.

BioSystems. 103, 1–12.
Gánti, T., 1971. Az élet princípium (The principles of life), first ed. Gondolat, Budapest (In Hungarian).
Gánti, T., 1978. Az élet princípium (The principles of life), second (revised) ed. Gondolat, Budapest (In Hungarian).
Gánti, T., 1987. The principle of life, first ed. OMIKK, Budapest.
Gilbert, W.,1986. Origin of life: the RNA world. Nature. 319, 618.
Hall, N., 2004. The quest for the chemical roots of life. Chem. Commun. 11, 1247–1252.
Hanczyc, M.M., Mansy, S.S., Szostak, J.W., 2007. Mineral surface brane assambly. Origins Life Evol. B. 37, 67–82.
Hazen, R.M., Filley, T.R., Goodfriend, G.A., 2001. Selective adsorption of L- and D-amino acids on calcite: implications for biochemical homochirality. P. Natl. Acad. Sci. USA. 98, 5487–5490.
Horowitz, N.H. 1945. On the evolution of biochemical synthe-ses. P. Natl. Acad. Sci. USA. 31, 153–157.
Ivica, N.A., Obermayer, B., Campbell, G.W., Rajamani, S., Gerland, U., Chen, I.A., 2013. The paradox of dual roles in the RNA world: resolving the conflict between stable folding and templating ability. J. Mol. Evol. 77, 55–63.
Johnson, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., Bartel, D.P., 2001. RNA-catalyzed RNA polymerization: accurate and general RNA templated primer extension. Science. 292, 1319–1325.
Johnson, A.P., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Lazcano, A., Bada, J.L., 2008. The Miller volcanic spark discharge experiment. Science. 322, 404.
Joshi, P.C., Aldersley, M.F., Ferris, J.P., 2011. Homochiral selectivity in RNA synthesis: montmorillonite-catalyzed quaternary reactions of D, L-purine with D, L- pyrimidine nucleotides. Origins Life Evol. B. 41, 213–236.
Joyce, G.F., 2002. The antiquity of RNA-based evolution. Nature. 418, 214–221.
Károlyi, Gy., Scheuring, I., Czárán, T., 2002. Metabolic network dynamics in open chaotic flow. Chaos. 12, 460–469.
Khersonsky, O., Tawfik, D.S. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu. Rev. Biochem. 79, 471–505.
Khvorova, A., Kwak, Y.G., Tamkun, M., Majerfeld, I., Yarus, M., 1999. RNAs that bind and change the permeability of phospholipid membranes. P. Natl. Acad. Sci. USA. 96, 10649–10654.
von Kiedrowski, G., Sijbren, O., Piet, H., 2010. Welcome Home, Systems Chemists! J. Sys. Chem. 1, 1.
Kim, P.-J., Jeong, H., 2005. Spatio-temporal dynamics in the origin of genetic information. Phisyca D. 203, 88–99.
Könnyű, B., Czárán, T., Szathmáry, E., 2008. Prebiotic replicase evolution in a surface-bound metabolic system: parasites as a source of adaptive evolution. BMC Evol. Biol. 8, 267.
Könnyű, B., Czárán, T., 2011. The evolution of enzyme specificity in the metabolic replicator model of prebiotic evolution. PloS One. 6, e20931.
Könnyű, B., Czárán, T., 2013. Spatial aspects of prebiotic replicator coexistence and community stability in a surface-bound RNA world model. BMC Evol. Biol. 13, 204.
Könnyű, B. Czárán, T., 2014. Phenotype/genotype sequence complementarity and prebiotic replicator coexistence in the metabolically coupled replicator system. BMC evol. Biol. 14, 234.
Kun, Á., Santos, M., Szathmáry, E., 2005. Real ribozymes suggest a relaxed error threshold. Nature Genetics. 37, 1008–1011.
Landweber, L.F., Simon, P.J., Wagner, T.A., 1998. Ribozyme engineering and early evolution. BioScience. 48, 94–103.
LaRowe, D.E., Regnier, P. 2008. Termodynamic potetial for the abiotic synthesis of adenine,

cytosine, guanine, thymine, uracil, ribose and deoxyribose in hydrotermal systems. Origins Life Evol. B. 38, 383–397.
Lilley, D.M.J., 2003. The origins of RNA catalysis in ribozyme. Trends. Biochem. Sci. 28, 495–501.
Nielsen, P.E., 2009. Peptid nucleid acids as prebiotic and abiotic genetic material, In Rasmussen, S., Bedau, M.A., Chen, L., Deamer, D.W., Krakauer, D.C., Packard, N.H., Stadler P.F. (Eds.), Protocells: Bridging non-living and living matter. MTI Press, Cambridge, pp. 337–346
Niesert, U., Harnasch, D., Bresch, C., 1981. Origin of life between Scylla and Charybdis. J. Mol. Evol. 17, 348–353.
Niesert, U., 1987. How many genes to start with? A computer simulation about the origin of life. Origins Life Evol. B. 17, 155–169, .
Ma, W., Yu, C., Zhang, W., Hu, J., 2007. Nucleotide synthetase ribozymes may have emerged first in the RNA world. RNA. 13, 2012–2019.
Manrubia, S.C., Briones, C. 2007. Modular evolution and increase of functional complexity in replicating RNA molecules. RNA. 13, 97–107.
Martin, W., Russel, M.J., 2003. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautrotrophic prokaryotes and from prokaryotes to nucleated cells. Philos. T. Roy. Soc. B. 358, 59–85.
Maynard-Smith, J., 1979. Hypercycles and the origin of life. Nature. 280, 445–446.
Maynard-Smith, J., Szathmáry, E., 1995. The major transitions in evolution. Oxford University Press, Oxford.
Meléndez-Hevia, E., Montero-Gómez, N., Montero, F., 2008. From prebiotic chemistry to cellular metabolism–the chemical evolution of metabolism before Darwinian natural selection. J. Theor. Biol. 252, 505–519.
Meszéna, G., Gyllenberg, M., Pásztor, E., Metz, J.A.J., 2006. Competitive exclusion and limiting similarity: an unfied theory. Theor. Popul. Biol. 69, 68–87.
Miller, S.L., 1953. A production of amino acids under possible Earth conditions. Science. 117, 528–529.
Miyakawa, S., Yamanashi, H., Kobayashi, K., Cleaves, H.J., Miller, S.L., 2002. Prebiotic synthesis from CO at-mospheres: implications for the origins of life. P. Natl. Acad. Sci. USA. 99, 14628–14631.
Monnard, P.-A., Kanavarioti, A., and Deamer, D.W., 2003. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. J. Am. Chem. Soc. 125, 13734–13740.
Morowitz, H.J., Kostelnik, J.D., Yang, J., Cody, G.D., 2000. From the cover: the origin of intermediary metabolism. P. Natl. Acad. Sci. USA. 97, 7704–7708.
Orgel, L.E., 2000. Self-organizing biochemical cycles. P. Natl. Acad. Sci. USA. 97, 12503–12507.
Orgel, L.E., 2004. Prebiotic chemistry and the origin of the RNA world. Crit. Rev. Biochem. Mol. 39, 99–123.
Parsons, I., Lee, M.R., Smith, J.V., 1998. Biochemical evolution II: origin of life in tubular microstructureson weathered feldspar surfaces. P. Natl. Acad. Sci. USA. 95, 15173–15176.
Patzke, V., von Kiedrowski, G,. 2007. Self replicating systems. Arkivoc. 8, 293–310.
Powner, M.W., Gerland, B., Sutherland, J.D., 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature. 459, 239–242.
Pross, A., 2004. Causation and the origin of life. Metabolism or replication first? Origins Life Evol. B. 34, 307–321.
Rasmussen, S., Bedau, M.A., McCaskill, J.S., Packard, N.H., 2009. A roadmap to protocells. In Rasmussen, S., Bedau, M.A., Chen, L., Deamer, D.W., Krakauer, D.C., Packard, N.H., Stadler P.F. (Eds.), Protocells: Bridging non-living and living matter. MTI Press, Cambridge, pp. 71–100.
Robertson, M.P., Joyce, G.F., 2010. The origins of the RNA world. Cold Spring Harb. Perspect.

Biol. 2, 1–22.
Rohatgi, R., Bartel, D.P., Szostak, J.W., 1996. Nonenzymatic,template-directed ligation of oligoribonucleotides is highly regioselective for formation of 3’ - 5’ phosphodiester bond. J. Am. Chem. Soc. 118, 3340–3344.
Rushdi, A.I., Simoneit, B.R., 2001. Lipid formation by aqueous Fischer-Tropsch-type synthesis over a temperature range of 100 to 400 degrees C. Origins Life Evol. B. 31, 103–118.
Scheuring, I., Czárán, T., Szabó, P., Károlyi, G., Toroczkai, Z., 2003. Spatial models of prebiotic evolution: soup before pizza? Origins Life Evol. B. 33, 319–355.
Schultes, E.A., Bartel, D.P., 2000 One sequence, two ribozymes: implication for the emergence of new ribozyme fold. Science. 289, 448–452.
Segré, D., Ben-Eli, D., Deamer, D.W., Lancet, D., 2001. The lipid world. Origins Life Evol. B. 31, 119–145.
Szathmáry, E., 1992. Viral sex, levels of selection, and the origin of life. J. Theor. Biol. 159, 99-109.
Szathmáry, E., 2006. The origin of replicators and reproducers. Philos. T. Roy. Soc. B. 361, 1761–1776.
Szathmáry, E., 2007. Coevolution of metabolic networks and membranes: the scenario of progressive sequestration. Philos. Trans. R. Soc. London B. 362, 1781–1787.
Szathmáry, E., Demeter, L. 1987. Group selection of early replicators and the origin of life. J. Theor. Biol. 128, 463–486.
Szathmáry, E., Gladkih, I., 1989. Sub-exponential growth and coexistence of non-enzymatically replicating templates. J. Theor. Biol. 138, 55–58.
Szathmáry, E., Santos, M., Fernando, C., 2005. Evolutionary potential and requirements for minimal protocells. Top. Curr. Chem. 259, 167–211.
Szilágyi, A., Zachar, I., Szathmáry, E., 2013. Gause’s principle and the effect of resource partitioning on the dynamical coexistence of replicating templates. PLoS Comput. Biol. 9, e1003193.
Takeuchi, N., Hogeweg, P., 2007. Error-threshold exists in fitness landscapes with lethal mutants. BMC Evol. Biol. 7, 15–27.
Toffoli, T., Margolus, N., 1987. Cellular automata machines: a new environment for modelling. MIT Press, Cambridge.
Wächtershäuser, G., 1990. Evolution of the first metabolic cycles. P. Natl. Acad. Sci. USA. 87, 200–204.
Wilson, D.S., 1975. A Theory of group selection. P. Natl. Acad. Sci. USA., 72, 143–146.
Wochner, A., Attwater, J., Coulson, A., Holliger, P., 2011. Ribozyme-catalyzed transcription of an active ribozyme. Science. 332, 209–212.

Figure legends

Graphical abstract: Metabolically coupled replicator system. The metabolic replicator system with four autocatalytic metabolic replicators (I_i, i = 1, .., 4 within the circular arrows). *M* is the metabolic reaction network supported by the metabolic replicators as enzymes (solid lines) and producing monomers for their replication (dashed lines).

Figure 1. Parasites of the hypercycle. P_1 : selfish parasite; P_2: short-cut parasite. Based on (Scheuring et al. 2003).

Figure 2. A schematic representation of the retroevolution of metabolism. The evolution of metabolically active replicators k (pV_k, k = n, m, q, …) catalysing an increasingly complex network (here: chain) of metabolic reactions (solid arrows and coloured folded structures) to produce monomer *V.* Reactions are included in the metabolic network sequentially as monomers (*V*) and then monomer precursors (*A* and *B*) are depleted from the environment (right diagram) by the increasing replicator population. *n*, *m* and *q* are stoichiometric constants.

Figure 3. The MCRS concept and neighbourhood definitions of the spatially explicit MCRS model. Panel A: The metabolic replicator system with four autocatalytic metabolic replicators (I_i, i = 1, .., 4 within the circular arrows). *M* is the metabolic reaction network supported by the metabolic replicators as enzymes (solid lines) producing monomers for their own replication (dashed lines). **Panel B**: The relation of metabolic (I_i, where i = 1, .., 3) and parasitic (*P*) replicators to metabolism. Parasites consume monomers produced by the metabolic network but do not contribute to metabolism by catalytic activity. **Panel C**: Neighbourhood definitions on the non-structured surface of the spatially explicit model. *X* is an empty site of the cellular automaton lattice, I_i (i = 1, .., 4) are the metabolic replicators. Dark grey sites are the replication neighbourhood of the empty site (*von Neumann neighbourhood* in this case) and light grey sites constitute the metabolic neighbourhood of replicator I_1 (3*x*3 *Moore neighbourhood* in this case). (From (Könnyű and Czárán, 2013))

Figure 4. Structure of a real mineral surface and the model. Panel A SEM image of a resin cast of an etch-pit network near the surface of a weathered Shap alkali feldspar (scale bar 20 lm). The cast was made by impregnating the feldspar with Araldite resin under vacuum, curing, and dissolving away the feldspar in concentrated HF. The surface of the feldspar is off the bottom of the micrograph, and the image is of a pile of two-dimensional networks that have fallen over to lie on top of each other. Because the resin is flexible, parts of the networks are curved. The original etch-pits were developed on edge dislocations very nearly parallel to b (*horizontal*) and c (*vertical*) in the perthite contact plane close to 601 of the monoclinic feldspar (SEM picture and caption from Fig.2. of (Parsons et al. 1998)). **Panel B** $I_{1..4}$ are the metabolic replicators; *M* is metabolism. Solid arrows represent the flux of resources (raw materials) outside the pores chemically transformed inside the pores in nucleotides by the catalytic activity of replicators (ribozyme). White arrows mean the catalytic effect of metabolic replicators helping metabolism. *P* represents a parasitic replicator that uses the monomers supplied by metabolism, but it does not help producing them. (From (Branciamore et al. 2009))

Figure 5.: A typical run of MRM. Parameters: system size (number of metabolic replicators) n = 4; replicator mobility D = 4, size of metabolic neighbourhood: 5x5 (Moore); size of replication neighbourhood: von Neumann; replication parameters of replicators: k_1 = 3 (blue), k_2 = 5 (red), k_3 = 7 (green), k_4 = 9 (orange),

Figure 6. Coexistence of metabolic replicators as a function of replicator diffusion (*D*), metabolic (*h*) and replication (*r*) neighbourhood size. The panels of the figure differ in the number of diffusion steps per generation: **Panel A**: $D = 0$, **Panel B**: $D = 1$, **Panel C**: $D = 4$ and **Panel D**: $D = 100$. *x*- and *y*-axes are the sizes of metabolic neighbourhoods (*h*) and replication neighbourhoods (*r*), respectively (*N*: *von Neumann* neighbourhood; 3: *3x3*, 5: *5x5*, 7: *7x7*, 25: *25x25* and 37: *37x37 Moore* neighbourhoods). The grayscale shades correspond to average replicator densities (% occupied) on the whole grid at the end of the simulations (i.e., at $t = 1.000$). The numbers within the cells of the tables indicate coexistent/extinct replicate simulations out of five repetitions with the same parameter set and different pseudo-random number sequences. Based on (Könnyű and Czárán, 2013).

Figure 7. The effect of migration and pore size on total replicator density in the pore-model. Fixed parameters: resource input ($r = 2$) and system size ($n = 5$). From (Branciamore et al. 2009).

Figure 8. The maximum number of coexisting metabolic replicators as the function of replicator diffusion (*D*), metabolic (*h*) and replication (*r*) neighbourhood size. The panels of the figure differ in the number of diffusion steps per generation: **Panel A**: $D = 0$, **Panel B**: $D = 4$ and **Panel C**: $D = 100$ *x*- and *y*-axes are the sizes of metabolic neighbourhoods (*h*) and replication neighbourhoods (*r*) respectively (*N*: *von Neumann* neighbourhood; 3: *3x3*, 5: *5x5*, 7: *7x7*, 25: *25x25* and 37: *37x37 Moore* neighbourhoods). The number within a cell of the panel shows the maximum attainable system size (n_{max}) for the corresponding parameter set. Other parameters: $p_d = 0.2$, $C_e = 2.0$, $k_i = 3.0 + 2.0i$ ($i = 0, .., n_{max}$). From (Könnyű and Czárán, 2013). **Panel D**: Relationship between system size and minimal pore size necessary for coexistence in the pore model. The migration parameter was $d = 0.8$. From (Branciamore et al. 2009).

Figure 9.: MRM and parasite(s). Panel A: Parameters: system size (number of replicators): 3 + parasite (black); *D*: 4, size of metabolic neighbourhood: 3x3 (Moore); size of replication neighbourhood: von Neumann; replication parameters of metabolic replicators: $k_1 = 3$ (blue), $k_2 = 5$ (red), $k_3 = 7$ (green), and parasite $k_p = 9$ (black). **Panel B:** Parameters: system size (number of replicators): 4 metabolic and 4 parasite replicators; *D*: 4, size of metabolic neighbourhood: 5x5 (Moore); size of replication neighbourhood: von Neumann; replication parameters of replicators: $k_{1m} = 3.0$ (blue), $k_{1p} = 4.0$ (light grey), $k_{2m} = 5.0$ (red), $k_{2p} = 6.0$ (middle grey) , $k_{3m} = 7.0$ (green), $k_{3p} = 8.0$ (dark grey), $k_{4m} = 9.0$ (orange) and $k_{4p} = 10.0$ (black); subscripts *m* and *p* denote metabolic and parasite replicator types, respectively.

Figure 10. Typical runs of specially modified MRM. Panel A: The pheno/geno version of MRM. Parameters: system size (number of replicators): 4 phenotype and 4 genotype replicators; *D*: 4, size of metabolic neighbourhood: 3x3 (Moore); size of replication neighbourhood: 37x37 (Moore); replication parameters of replicators: $k_{1p} = 3.0$ (blue), $k_{1g} = 4.0$ (blue), $k_{2p} = 5.0$ (red), $k_{2g} = 6.0$ (red) , $k_{3p} = 7.0$ (green), $k_{3g} = 8.0$ (green), $k_{4p} = 9.0$ (orange) and $k_{4g} = 10.0$ (orange); subscripts *p* and *g* denote phenotype-forms (solid lines) and genotype-forms (dashed lines) of replicator types, respectively. **Panel B:** MRM with a facultative metabolic cooperator. Parameters: system size (number of replicators): 3 essential metabolic and facultative metabolic replicators; *D*: 4, size of metabolic neighbourhood: 3x3 (Moore); size of replication neighbourhood: von Neumann; replication parameters of replicators: $k_1 = 3$ (blue), $k_2 = 5$ (red), $k_3 = 7$ (green), and the facultative cooperator: $k_p = 9$ (orange).

Figure 11. The $E_1 - E_2 - k$ trade-off surface. The trade-off function constrains the phenotypes of emerging mutant replicators to below the surface given by

$$k(E_1, E_2) = \left[E_{\max}^{g} - \left[E_1^{b} + E_2^{b} \right]^{\frac{1}{b}g} \right]^{\frac{1}{g}} \cdot \frac{k_{\max} - k_{\min}}{E_{\max}} + k_{\min}.$$

Fixed parameters: k_{min} = 2.0, k_{max} = 4.0, E_{max} = 10.0.

Panel A: convex function representing strong trade-off both between the two enzyme activities E_1/E_2 and between enzyme activities and replication rate, *E/k*. **Panel B:** a function with convex (strong) E_1/E_2 trade-off and concave (weak) *E/k* trade-off . **Panel C:** concave (weak) E_1/E_2 and convex (strong) *E/k* trade-off. **Panel D:** both the E_1/E_2 and the *E/k* trade-offs are concave (weak). From (Könnyű and Czárán, 2011).

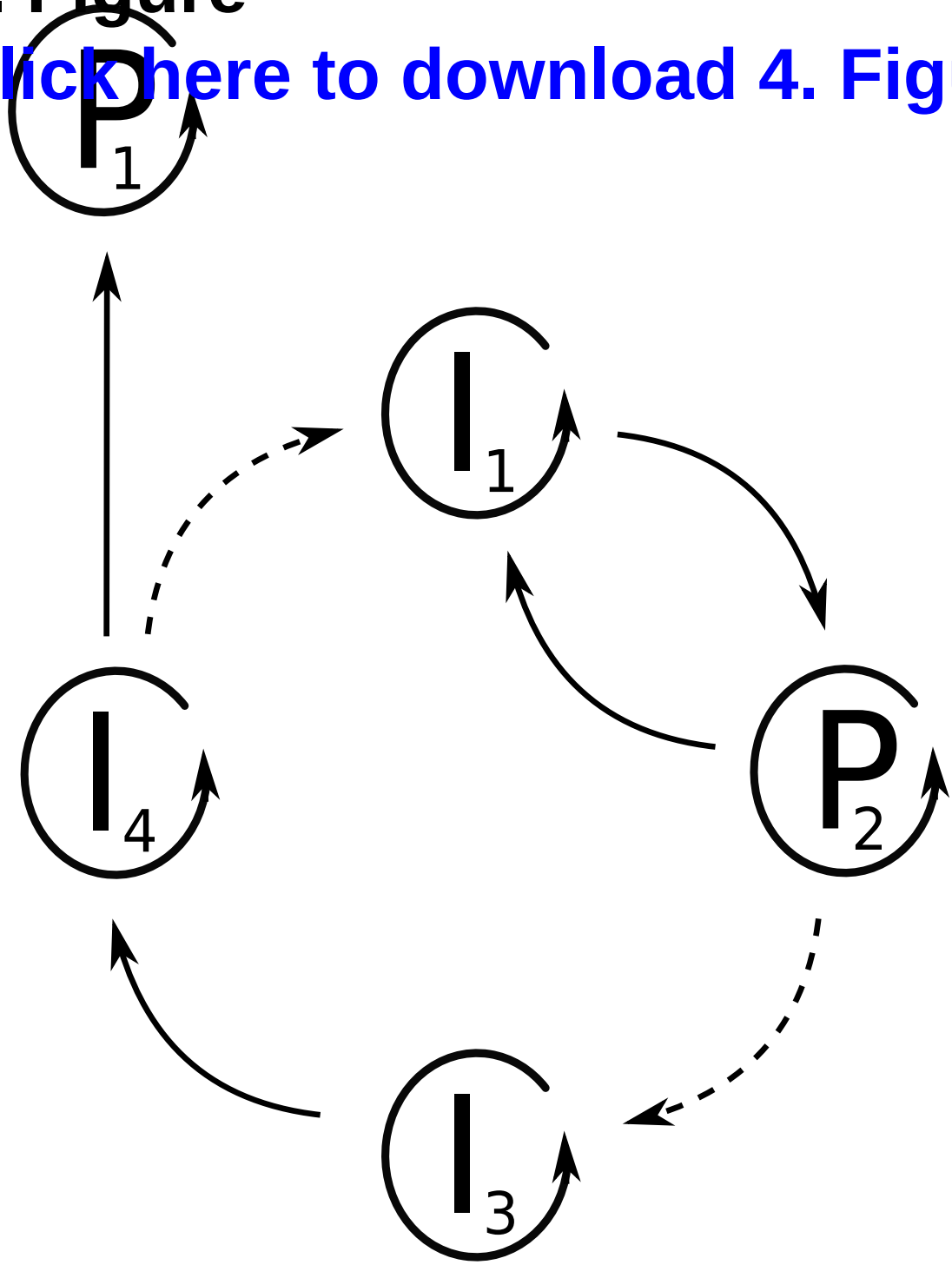
Figure 12. Frequencies of replicator types. **Panel A**: The steady-state frequencies of specialist and generalist replicators as a function of *b* (the strength of the trade-off between enzymatic activities), at D = 0; **Panel B**: the same, at D = 5. Other parameters: p_m = 0.01 (mutation rate), g = 1.0 (the strength of the trade-off enzymatic activities and replication rate), E_{max} = 10 (maximal enzymatic activities) and k_{max} = 2.5 (maximal replication rate) at the 150.000th generation. Note that the frequency of parasitic replicators is less than 1% everywhere in this parameter setting, so we have not plotted it here. Based on (Könnyű and Czárán, 2011).

Figure 13. The benefit of evolving a sequence-aspecific replicase replicator. Panel A: Metabolic system with a parasite evolved into a replicase (*R*). Dashed-dotted lines represent sequence-aspecific replicase activity. Other arrows and letters are the same as in Figure 1. **Panel B:** The effect of an evolving replicase replicator on the dynamics of the metabolic replicator community. Replication parameters: k_1 = 2 (blue), k_2 = 4 (red), k_3 = 6 (green), and parasite/replicase k_p = 8 (orange). Black line: replicase activity (scale on the second *y* axis). Based on Könnyű et al. 2008.

Figure 14. The Metabolically Coupled Repricator System enclosed in a self-produced membrane vesicle ("protocell"). The metabolic replicator set ($I_{1..4}$) with a replicase (***R***), a lipid synthetase (***L***) and a membrane channel forming replicator (***T***) added. **M** produces membranogenic molecules (*black triangles*) which are transformed to membrane molecules (black rectangles) by the lipid synthetase (***L***) replicator. New lipid molecules are inserted into the membrane spontaneously. Transporter replicators (*grey rectangle with a T*) insert themselves into the membrane to form trans-membrane channels which selectively let small metabolic precursor molecules (*black stars)* enter the vesicle.

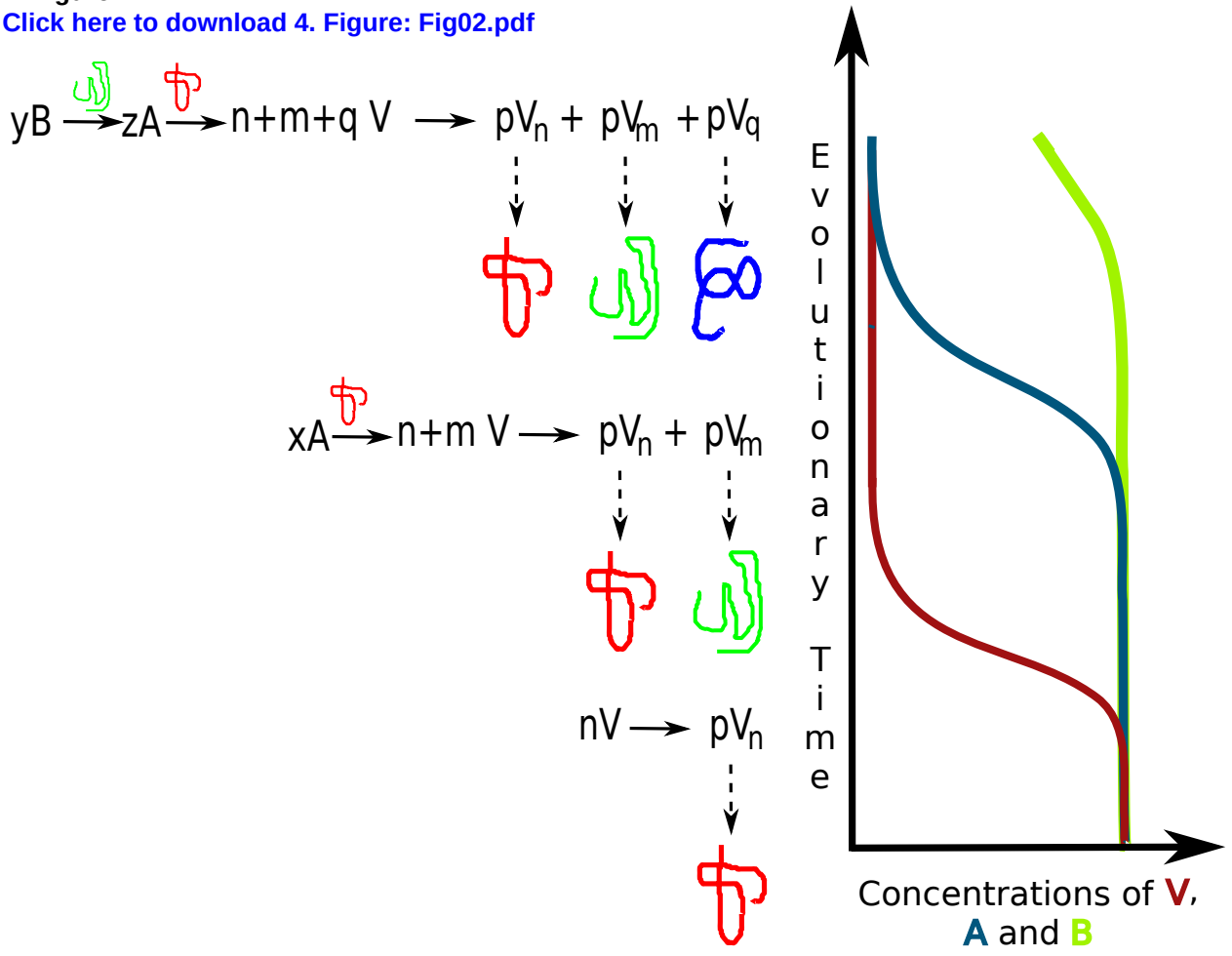
4. Figure

Click here to download 4. Figure



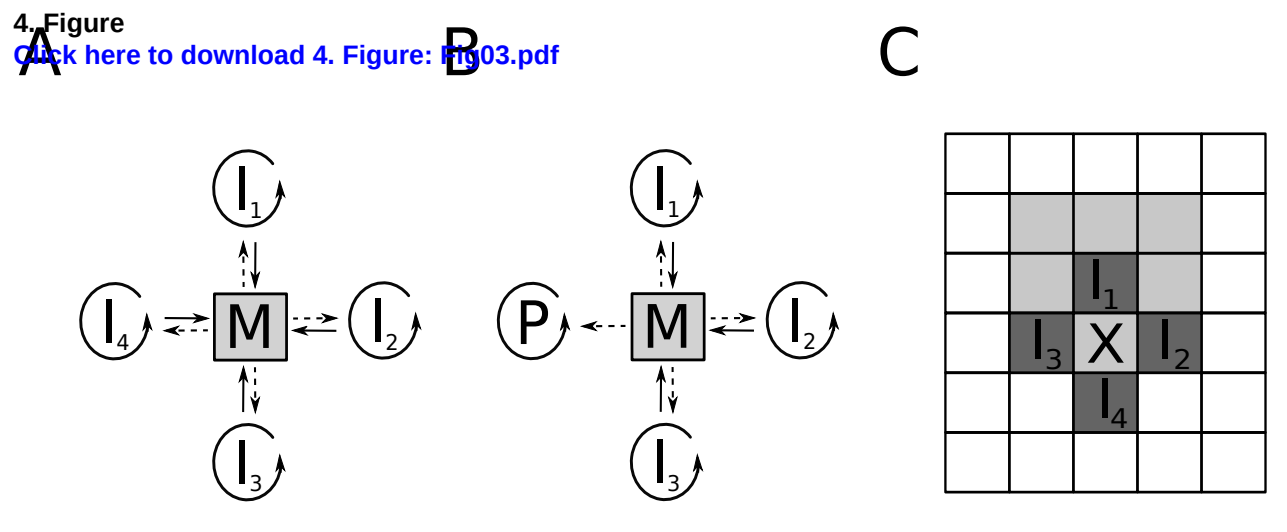
4. Figure

Click here to download 4. Figure: Fig02.pdf



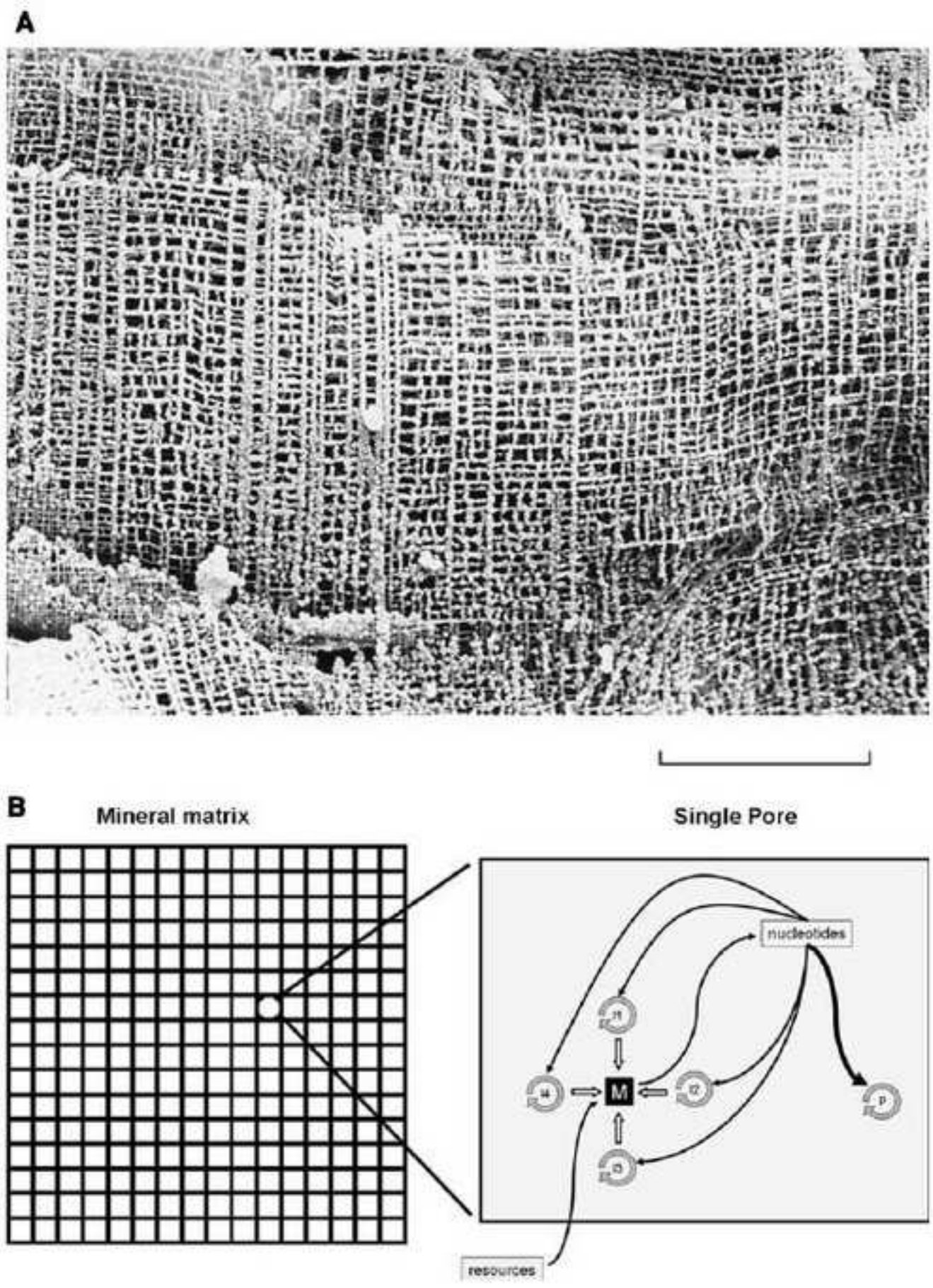
4. Figure

Click here to download 4. Figure: Fig03.pdf



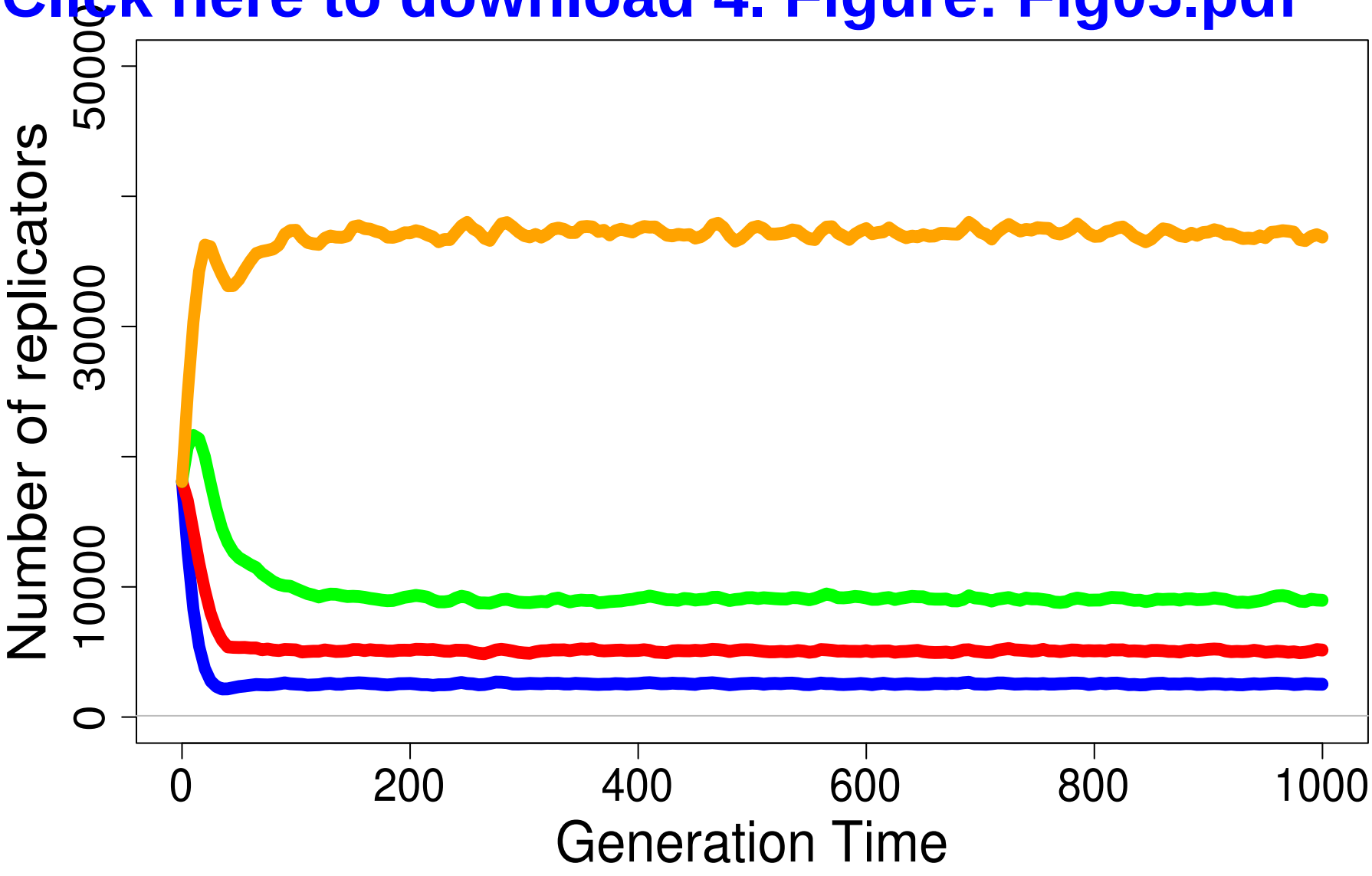
4. Figure

Click here to download high resolution image



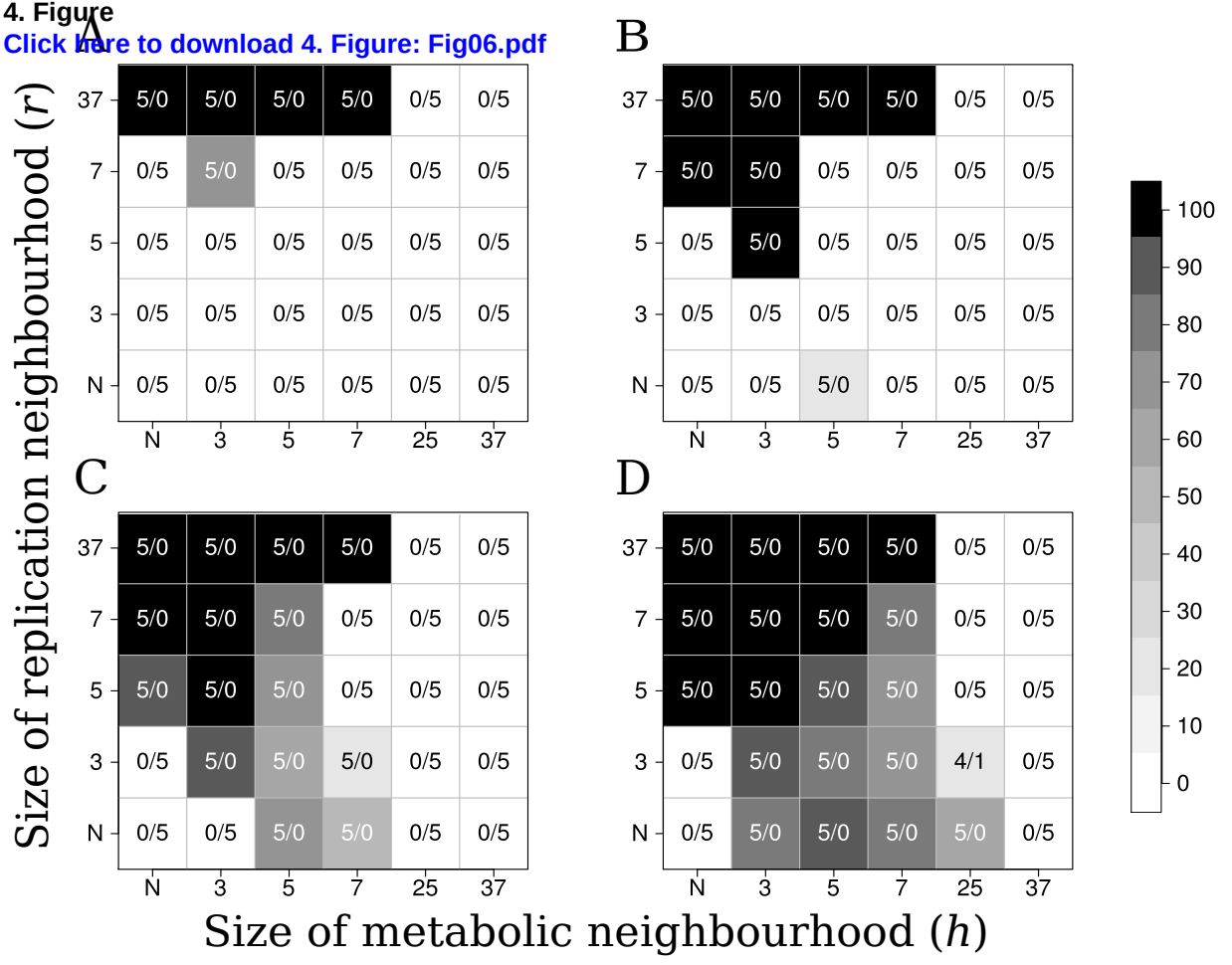
4. Figure

Click here to download 4. Figure: Fig05.pdf



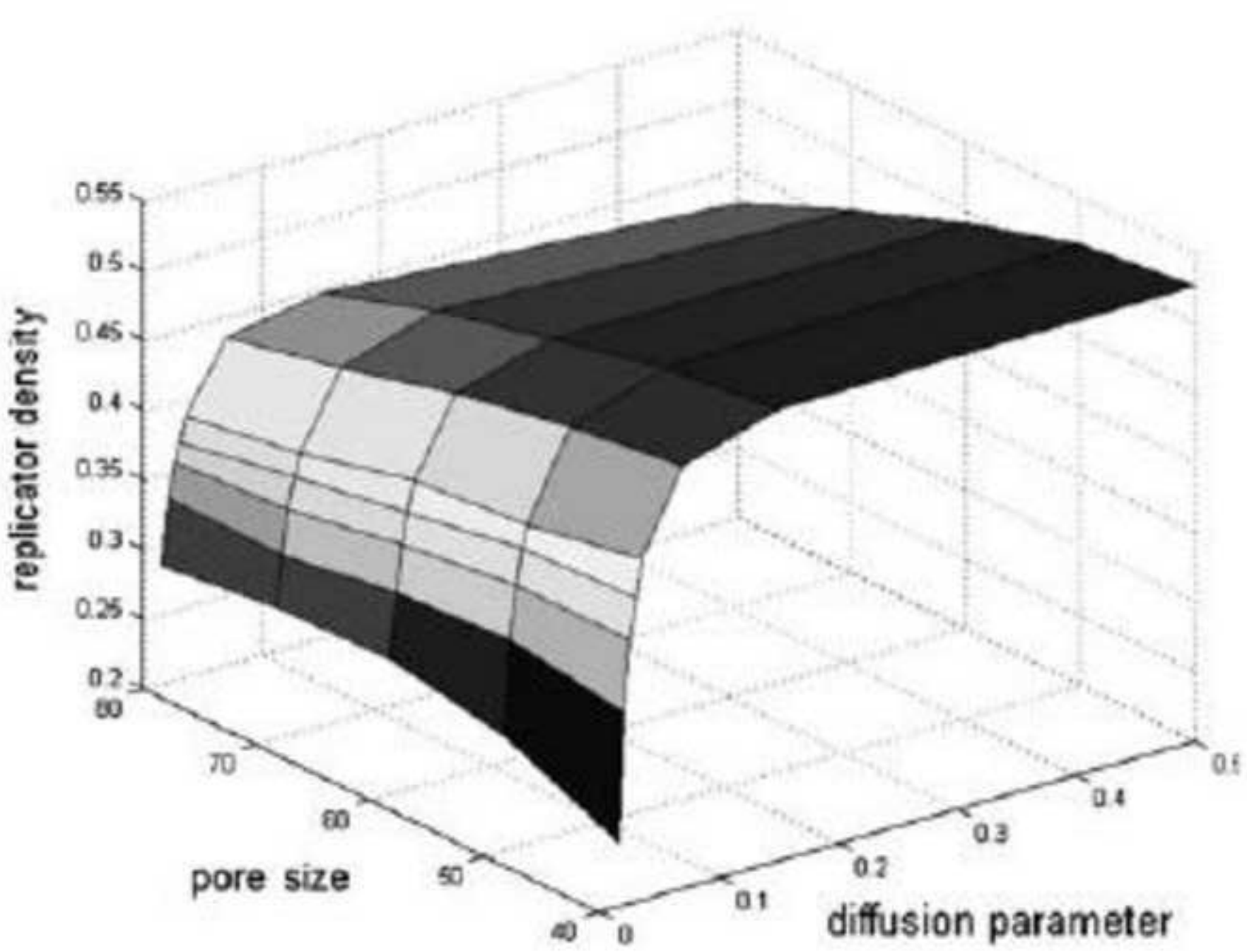
4. Figure

Click here to download 4. Figure: Fig06.pdf



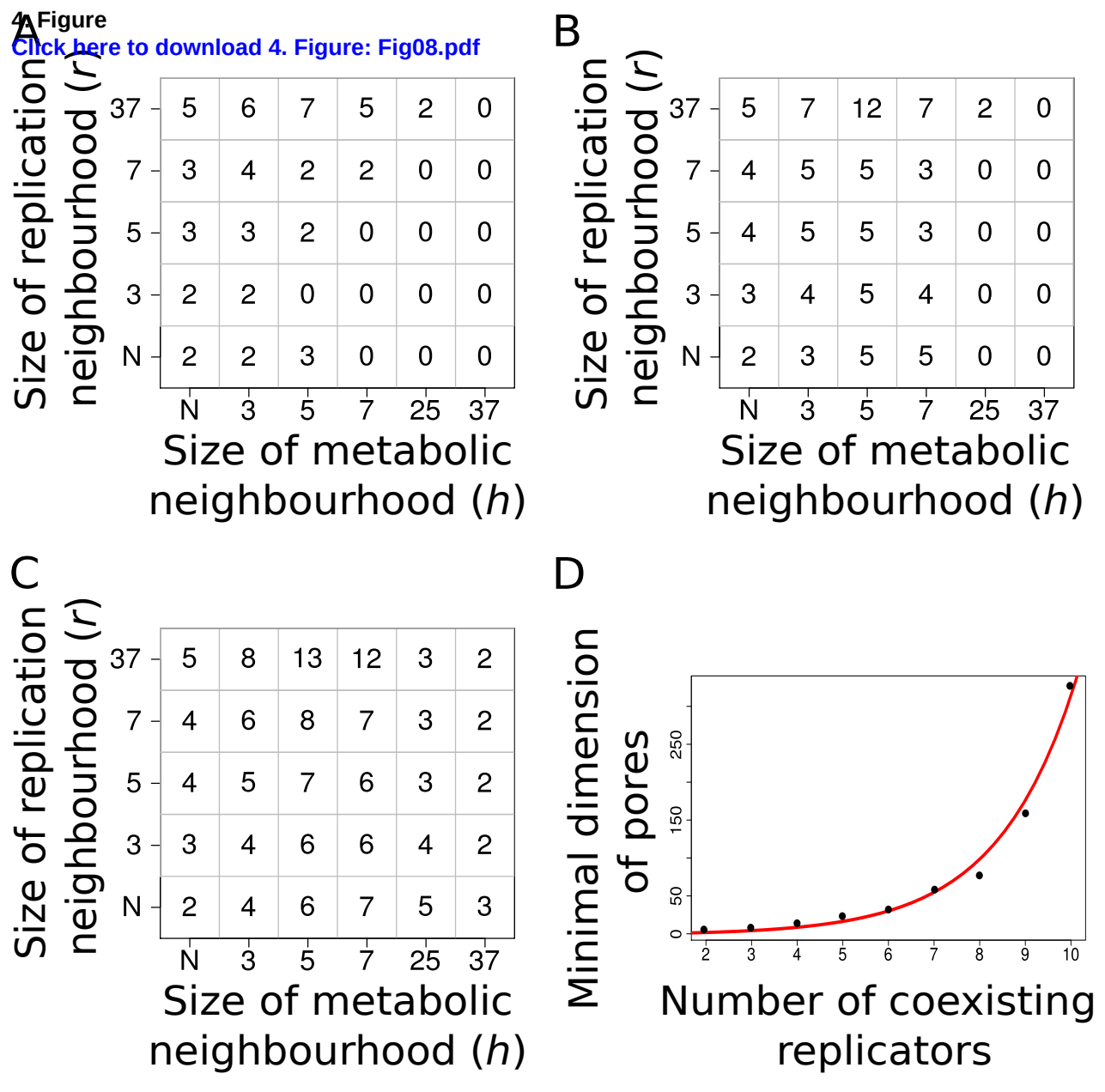
4. Figure

Click here to download high resolution image



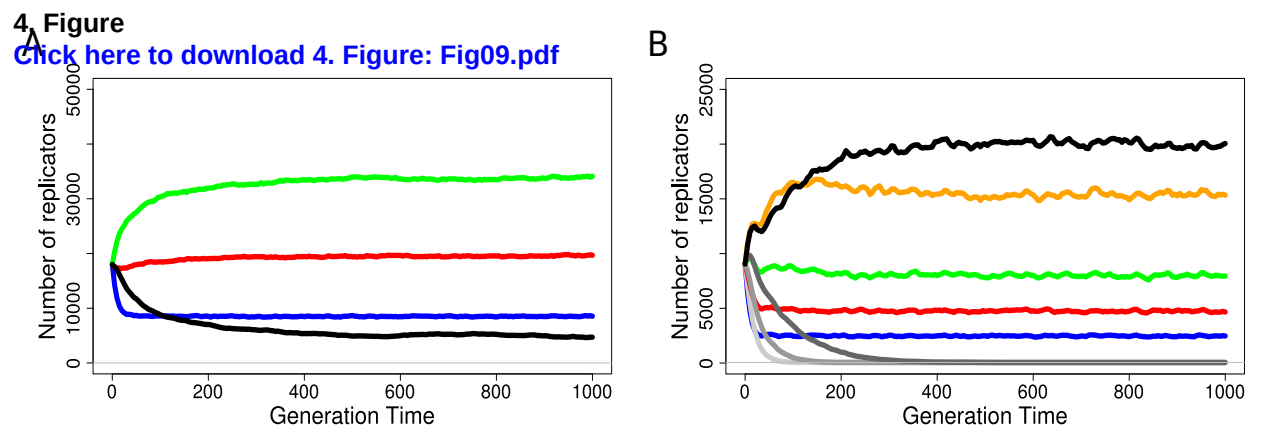
4 Figure

Click here to download 4. Figure: Fig08.pdf



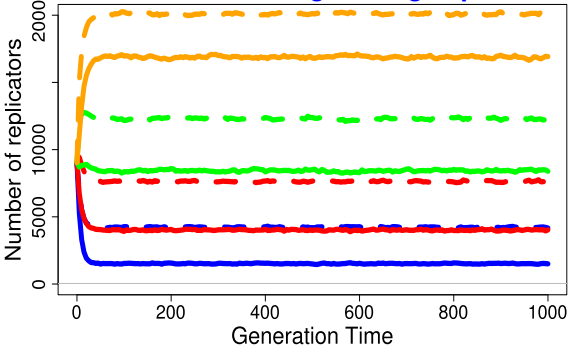
4. Figure

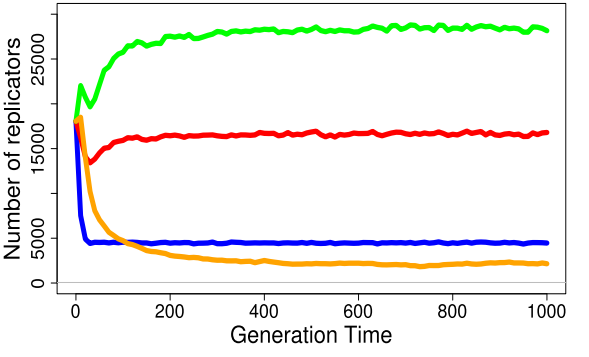
Click here to download 4. Figure: Fig09.pdf



4 Figure

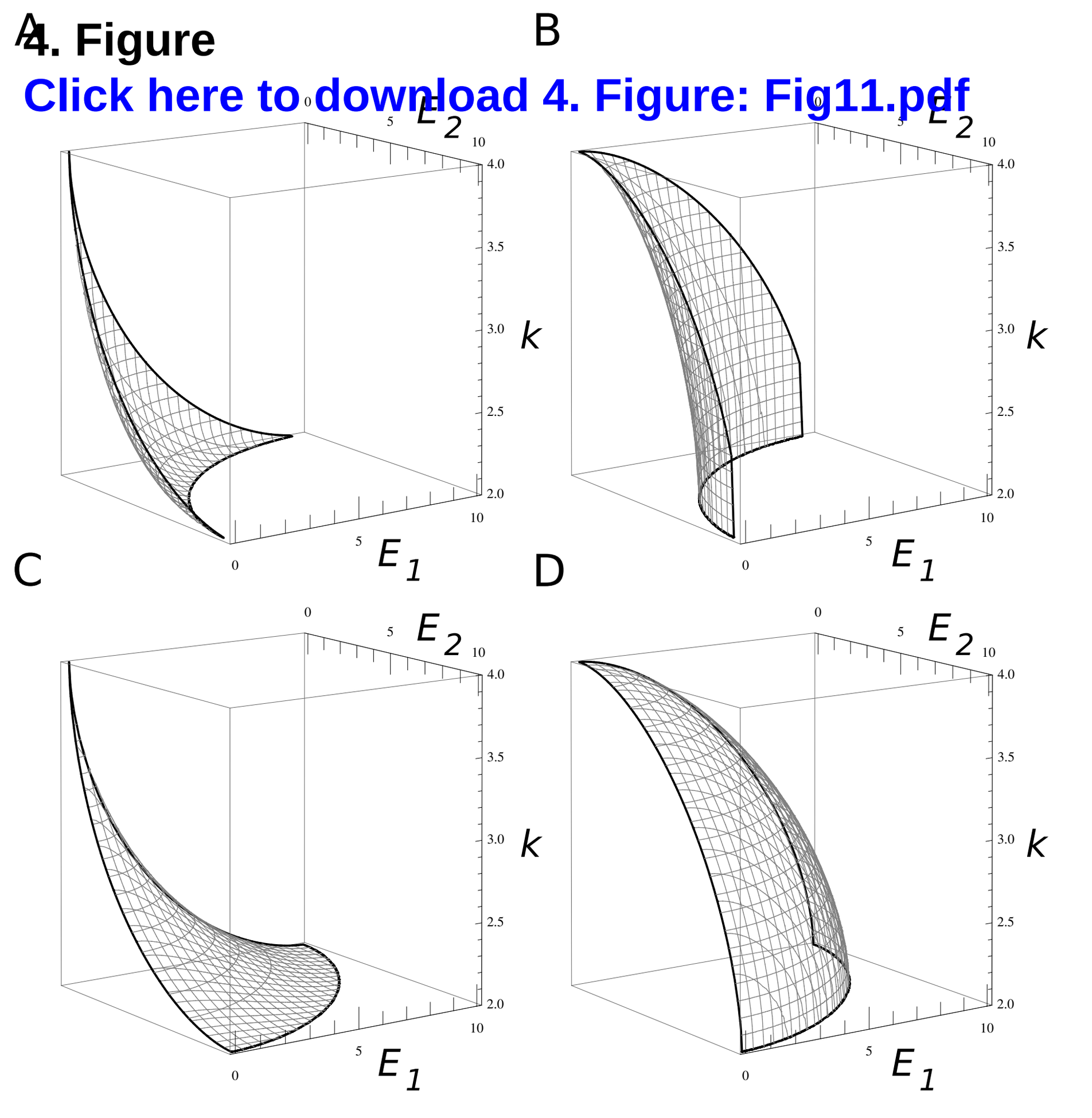
Click here to download 4. Figure: Fig10.pdf





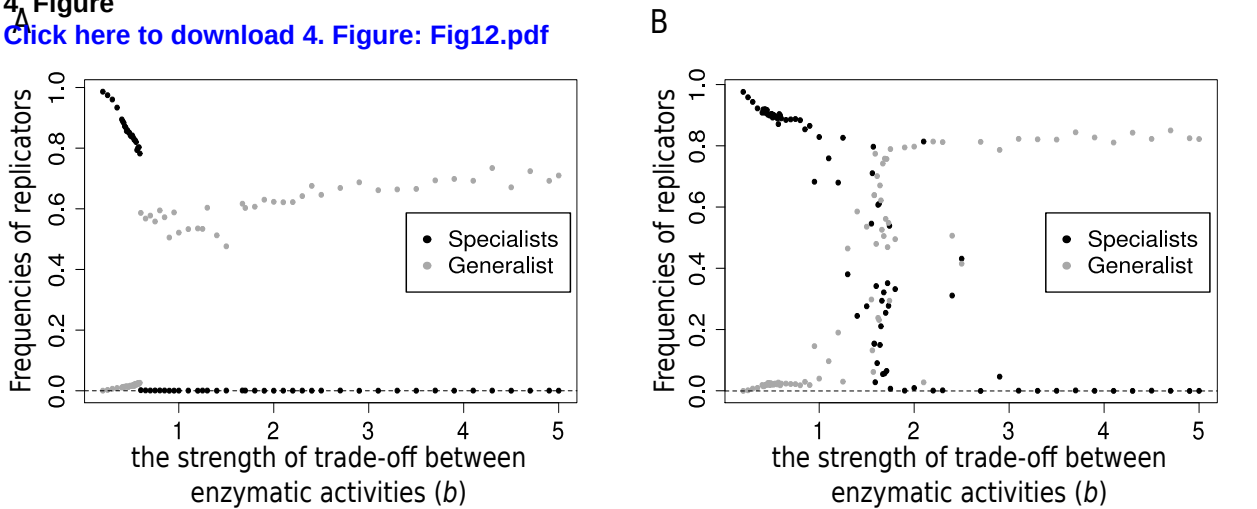
4. Figure

Click here to download 4. Figure: Fig11.pdf



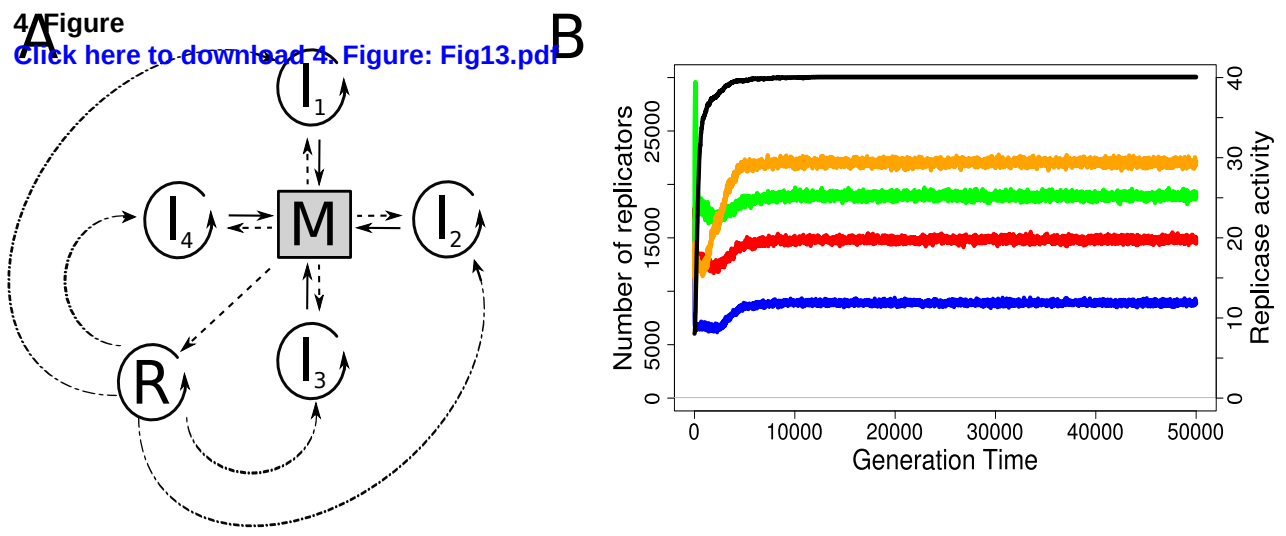
4. Figure

Click here to download 4. Figure: Fig12.pdf



4 Figure

Click here to download 4. Figure: Fig13.pdf



4. Figure

Click here to download 4. Figure: Fig14.pdf