Biomathematical model of prebiotic autocatalytic RNA networks shows birth of adaptive immunity from degenerating molecular parasite catalysts

Short title:

degenerating molecular parasites seed immunity

Authors:

Magnus Pirovino[1], Christian Iseli[2], Joseph A. Curran[3], Bernard Conrad[4*]

Affiliations: [1]OPIRO Consulting Ltd., FL-9495 Triesen, Principality of Liechtenstein, [2]Bioinformatics Competence Center, EPFL and Unil, CH-1015 Lausanne, [3]Department of Microbiology and Molecular Medicine, and Institute of Genetics and Genomics of Geneva (iGE3), Geneva University, CH-1211 Genève 4, [4]Genesupport, Avenue de Sévelin 18, CH-1004 Lausanne

*Correspondence to: bernard.conrad@genesupport.ch

Catalysis and specifically autocatalysis are the quintessential building blocks of life. However, autocatalytic networks are necessary but not sufficient for life-like properties such as self-propagation and adaptation to emerge. We previously showed using Lotka-Volterra equations that a tripartite, early life cycle comprising a host-template, molecular parasites and hyperparasites (parasites of parasites) conferred superior stability to the habitat. Here we build on this model, and on the seminal notion that chemical networks and ecologically interacting biological species are akin, implying an instructive equational analogy between catalysis and ecology, both accurately modelled by equations from the Lotka-Volterra family. Template-based autocatalytic RNA networks seed the spontaneous emergence of immunity against molecular parasites, with degenerating molecular parasite catalysts being the founding seed of the homeostatic process. First, upon molecular parasite encounter, specific ribocatalysts providing parasite resistance appear in primordial autocatalytic cycles, supplanted by promiscuous, efficiently self-replicating, but parasite sensitive ribozymes. This is supported by catalyst tradeoff analysis: substrate promiscuity confers high catalytic activity, but entails parasite exposure, higher substrate specificity implies relatively lower activity, yet it offers parasite resistance. Second, sustained parasite influx generates hyperparasite cycles, memorizing parasite-specific catalyst-subunits that now once more provide parasite resistance, in a homeostatically stabilized habitat. Third, under continuous exposure to progressively elaborate parasite populations, hyperparasite catalysts degenerate and embody antiparasite immunity. The observed triggering of microbial immunity by hyperparasitic microbial genomes are in agreement with this model. As such, it offers an attractive and unifying theory for the spontaneous birth of immunity in early life.

Author Summary

The quintessential components of life comprise a potent acceleration of naturally occurring, but improbable chemical reactions (catalysis), and the arrangement of accelerated chemical reactions in closed loops (autocatalytic sets). This is required, but is not sufficient for such networks to self-propagate and adapt. We previously developed an early life model for a stabilized habitat of networks, using a family of equations shown to accurately model both chemical reactions and the interactions of biological species. This model became stabilized only if the molecular host species was parasitized, an event considered unavoidable, and at the same time, the parasite itself was also parasitized. We now built on that model, using the same family of equations, by letting individual autocatalytic sets interact. The new results show that primitive autocatalytic networks emerge spontaneously, but inefficiently and resist parasites. With time, they become more efficient, which

renders them parasite-sensitive. Upon continuous external parasite invasion, the primitive, parasite-resistant catalysts (hyperparasites) that still persist are recruited, conferring stability to the habitat. If the habitat is further overwhelmed with parasites, it selects the elegant solution of letting the parasite-resistant catalysts degenerate, which now produces immunity to those parasites ("antibodies"). This represents the birth of adaptive immunity.

Introduction

The multistep transition process from inanimate to living systems has been dated to >3.9 Ga [1]. A non-exhaustive list of thresholds that needed to be progressively surmounted during this process has been drawn, comprising, among others, chirality symmetry breaking, spontaneous polymerization, self-assembly of compartments, and particularly the catalytic closure threshold [2]. While autocatalysis is the core ingredient of life, it is not sufficient. The ability of prebiotic networks to self-propagate and adapt hinges upon multiple, sequential interactions of autocatalytic subsystems among themselves and with the environment, creating higher order structures [3-7]. In one of those models, pairs of autocatalytic cycles following chemical mass action kinetics [8] exhibit the superior complexity-building behaviors usually attributed to biological species, for instance competitive, predator-prey or mutualistic relationships [9]. Predator-prey type autocatalytic cycles display a dynamics similar to the Lotka-Volterra (LV) model [9].

One central pillar in the theory of how early life could have materialized pertains to the "RNA world". This stipulates that the initial informational polymer was RNA, and that the same molecule was also the fundamental catalyst [10]. The underlying concept of autocatalytic sets was laid out more than half a century ago [11,12]. It was subsequently refined [3] and validated in a proof of principle experiment showing that RNA fragments could build self-replicating ribozymes from individual RNA fragments via cooperative catalytic cycles [13]. In parallel, an RNA polymerase ribozyme was generated by *in vitro* evolution that exhibited a high level of processivity [14]. Collectively, both theoretical considerations and experimental evidence support the view of a gradual evolution towards RNA self-catalysis. This started with self-assembly, followed by the formation of autocatalytic networks using self-assembled substrates. From this emerged different ribozymes (recombinases, ligases) and template based ligation, leading to a template based, self-replicase catalyzed assembly of monomers [15].

RNA self-catalysis is unthinkable without access to generic energy levels. It is commonly argued that autocatalytic chemical networks preceded template-based autocatalysis [16]. Indeed, proto-metabolic reactions without metals and enzymes have been demonstrated, suggesting the

existence of an ancestral analogue of the core reverse Krebs cycle [17]. By force of circumstances, autocatalytic metabolic and template-based networks are bound to have joined forces to render the emergence of protocells sufficiently likely [18].

The ribozyme catalyzed RNA-polymerization process is inherently error prone [10,14], generating variant sequence space mutants called a quasispecies [19,20]. This variant sequence-space intrinsically implies the obligate presence and evolutionary persistence of the molecular parasites associated with all life forms, as argued essentially for thermodynamic reasons [21,22]. We previously developed a prebiotic hyperparasite framework based on the LV-equations [23], demonstrating that superior homeostasis arose from the tripartite configuration within a given molecular habitat, namely, the molecular host, the parasite and the hyperparasite, i.e. the parasite of the parasite. The empirical evidence of highly prevalent tripartite communities in the virosphere lends substantial credit to this notion [24-27]. The LV-equations have become the default tool to quantitatively assess interactions in microbial communities [28]. However, it is often forgotten that Alfred Lotka initially modelled catalysis and autocatalysis [29]. Based on his equations modeling chemical reactions that transform a substrate [S] into a product [P] via two intermediates, X and Y, if the reactions producing X and Y are assumed to be autocatalytic, then the resulting ordinary differential equation is the classical LV-predator-prey system [30]. This stoichiometry corresponds to the Michaelis and Menten mechanism - the universal formalism for catalysis - that considers fully reversible reactions $S + E \rightleftharpoons B1 \rightleftharpoons B2 \rightleftharpoons P + E$; ($B1$ and $B2$ are the intermediates equivalent to X and Y) [31]. In fact, LV-equations commonly used in ecology, the Rosenzweig-MacArthur model with Holling type II response [32,33], are analogous to the Michaelis-Menten mechanism in enzyme kinetics [34]. Jointly, these considerations lay the ground for modelling catalysis, autocatalysis and interactions within primordial, template-based ecological RNA communities with one single biomathematical tool, namely the LV-model. This reasoning has led us to extend the initial tripartite model, building on multiple, sequentially interacting autocatalytic RNA networks memorizing the historical parasite exposure. We model, how molecular parasite resistance and antiparasite immunity can be selected for and gradually emerge from progressively degenerating hyperparasites in an intrinsic fashion in early life, in hosts exposed to increasingly more complexed parasite populations.

Results

In a stepwise and gradual process described in the Supplementary Information (proofs, Suppl. Inf.) and Methods sections, we lay out using the LV-equations, how simple template-based RNA autocatalytic cycles are initially built and how over time they become increasingly more elaborated as a result of molecular parasite exposure. Metabolic autocatalytic networks are assumed to be present

132 and to provide sufficient energy [23]. The coexistence of prebiotic ecology/heredity [35,36] with an
133 emerging canonical, template-based Darwinian heredity allows for the complexification of the
134 quasispecies that in turn reflects its historical parasite exposure. The previously established tripartite
135 habitat framework featuring a template-based RNA host, a molecular parasite that is itself parasitized
136 (hyperparasite; [23]), is instrumental for homeostasis and the emergence of resistance, and finally
137 immunity to the increasingly more exhaustive molecular parasite populations.
138
139 ***Autocatalytic cycles in transitional early RNA-settings***
140 The seminal duality of the RNA world, namely the fact that the same RNA molecule provides both the
141 template and the self-replicating catalyst [10] constitutes the foundation of the initial autocatalytic
142 cycles (replication equations (1.1), (1.2), forming a two-step autocatalytic cycle; proof outlined in
143 Suppl. Inf.). Since the template and the enzyme/ribozyme remain in principle unchanged during this
144 process, both can be viewed as catalysts, as long as the molecules do not undergo deleterious loss of
145 function mutations.
146
147 ***Molecular parasitism in autocatalytic cycles***
148 Molecular parasites are an inherent feature of the replication process, they are inseparably
149 associated with all life forms [21,22], for a number of reasons, the most important being the
150 accuracy-rate tradeoff [37,38]. Average enzymes operate orders of magnitude (k_{cat}/K_M of $\sim 10^5\ s^{-1}M^{-1}$)
151 below the diffusion rate (10^8-$10^9\ s^{-1}M^{-1}$), which implies promiscuous substrate binding, high catalytic
152 activity, and as a secondary tradeoff a high mutation rate with a large variance in the quasispecies
153 (speed being more important than enzyme accuracy; [39]). In an initial autocatalytic cycle, self-
154 replicating ribozymes emerge (ribozyme R_2 acting on a replicating template R_1; proof outline in
155 Suppl. Inf.; Model $1^{(1)}$ in Methods). The catalyst tradeoff opposes specific, but catalytically relatively
156 less active molecules, and promiscuous, yet catalytically highly efficient molecules [37,38]. Therefore,
157 the catalytically very efficient and promiscuous R_2 will inevitably be parasitized by RNA templates P.
158 Figure 1 illustrates the population sizes of the habitat constituents under high (Fig. 1A), and low (Fig.
159 1B) parasite exposure. As expected, high parasite loads overwhelm the habitat with P templates, an
160 observation that begs a solution.
161
162 ***A stable, primordial autocatalytic cycle in the presence of molecular parasites***
163 In a primordial habitat with abundant inflow of molecular RNA species including parasites, a
164 quasispecies will eventually emerge with the central feature that it also comprises molecules with
165 superior parasite-specificity, which implies relatively lower catalytic activity, and confers parasite
166 resistance (protoreplicase E_2, prototemplate F_1; equations and proofs outlined in the Suppl. Inf.).

This composite feature progressively arises in successive rounds of autocatalytic cycles, initially comprising the encounter of the catalyst $E_{1,}$ provided by a neighboring habitat, with F_1 evolved autocatalytically in the habitat, producing the primordial catalyst $E_{2,}$ that is capable of amplifying F_1 but not P. It is the twofold consequence of the primordial parasite-specificity of $E_{2,}$, and of the complexity of P that exceeds the processivity of the protoreplicase. As a corollary and as displayed in Fig. 2, selection of such molecules confers considerable resistance to parasite invasion, and restores host template and catalyst abundance (Model $1^{(0)}$, Methods). There is both, strong theoretical and experimental evidence to support these findings. Indeed, while it is generally assumed that catalysts arising in early autocatalytic cycles have low activity and high promiscuity [37], an in depth physico-chemical analysis of how enzymes meet demands of selectivity and accuracy points towards accuracy-rate tradeoffs [38]. Selectivity (accuracy, specificity) implies the ability to discriminate between two templates, provided both are present [37,38], and rate denotes speed [39]. Ground state discrimination means that specificity is achieved mainly through substrate binding, which imposes strong accuracy-rate tradeoffs [38]. Improvements in selectivity mediated by tighter cognate substrate binding invariably lead to lower catalytic efficiency (parallel decreases in the constants for the cognate substrates K_M^{cog} and k_{cat}^{cog}; [38]). This is precisely what the model stipulates, the protoreplicase has a primal parasite-selectivity, which comes at the cost of a relatively reduced catalytic activity, yet it allows for adequate host molecule-, but not parasite-amplification. In conclusion, if catalysts with relatively reduced function in the quasispecies confer parasite resistance, yet restore host catalyst and template abundance, they will be positively selected [35].

Evolution of the primordial autocatalytic cycle into a full-fledged autocatalytic cycle

Natural selection in the mutant quasispecies space will ultimately produce, starting from E_2, a fully efficient, self-replicating ribozyme (replicase R_2) that is bestowed with high catalytic activity [14], capable of generating all enzymes and templates in the habitat (equations and proofs outlined in the Suppl. Inf.). However, this has a cost. R_2 is promiscuous, catalytically highly efficient and therefore exposed to parasite invasion, which was not the case for its primordial catalyst precursor with relatively lower activity (protoreplicase E_2; Model $1^{(1+)}$ stabilized host-parasite-hyperparasite cycles, Methods).

At this point, the protoreplicase E_2-subdomain of R_2 confers P selectivity, restoring parasite resistance and rescuing the abundance of the partial catalyst and its templates, even with a moderate parasite inflow ($\alpha > 0$, Fig. 3 left panel). This process has as its central feature a hyperparasite cycle (equations and proofs outlined in the Suppl. Inf.). It is further noteworthy that F_2 (formerly E_2) is still produced (Fig. 3), and is acting additively to R_2, which is important to generate antiparasitism later.

Under permanent and moderate, thus manageable parasite inflow ($P^{special}$, Suppl. Inf.), an increasingly ramified network of autocatalytic cycles arrives in the habitat that persists even when the parasite exposure ceases, constituting a historical record of past parasite exposures ($\alpha = 0$, Fig. 3 right panel).

There is considerable evidence in favor of the essential components of this model, in particular the catalyst accuracy-rate tradeoff and the hyperparasite cycle. Firstly, a solid body of theoretical and experimental evidence supports catalyst accuracy-rate tradeoffs [37,38]. Secondly, metagenomic studies demonstrate the virtually universal presence of tripartite constituents across molecular parasite habitat classes, most persuasively for bacteria-phages-phage satellites [24,25], for RNA viruses partnering the host- virus-defective interfering RNAs [26], and for NCLDVs, with eukaryotic host-NCLDV-virophage [27]. Thirdly, additional support comes from the observation that hyperparasites are induced across the range of parasite concentrations, and especially at high multiplicity infection for RNA viruses [40,41].

Emergence of antiparasitism in autocatalytic cycles

Under conditions, where the sheer number and complexity of incoming molecular parasite populations would overwhelm the carrying capacity of the habitat, a degenerative process is initiated that erodes catalysis. Formally, the hyperparasite cycle degenerates to an antiparasite cycle (equations and proofs outlined in the Suppl. Inf.; Model $1^{(1-)}$, Methods). In Figures 4 and 5, the transition from hyperparasitism to antiparasitism is illustrated, in the sense that both coexist at equal activity, partial catalyst and non-catalytic antiparasitic activities (Fig. 4A, see Methods for parameters). In Figure 4B, antiparasitism is fully implemented; hyperparasitism has been fully replaced by the non-catalytic antiparasitism (reproduction parameters hyperparasite d and δ_3 are set to zero: $\delta_3 = d = 0$), making the difference between pure antiparasitism, and the transition from hyperparasitism to antiparasitism minor. In Figure 4B, the population fluctuations are more volatile than in Figure 4A, but pure antiparasitism does effectively stabilize the autocatalytic cycle, without a hyperparasite component. Figure 5 shows the gradual transition from pure hyperparasitism (top panel) to antiparastism (botton panel), in four successive steps (0, 0.25, 05. 0.75 and 1), progressively replacing hyperparasitism with antiparasitism. Therefore, we conclude that antiparasitism can replace hyperparasitism in situations with a stronger and highly diverse parasite inflow that would overrun the carrying capacity of the autocatalytic cycle in the habitat.

Discussion

235 We have outlined in a stepwise fashion, how increasingly elaborate autocatalytic cycles gradually
236 emerge in an early life RNA-habitat. Initially, catalysts with low catalytic activity are selected, because
237 they are parasite-specific and confer resistance (primordial autocatalytic cycles). Soon, complexed
238 mixtures evolve, featuring self-replicating catalysts of high catalytic activity that become stabilized
239 through the selection and integration of hyperparasite cycles, including relatively less active catalyst-
240 subunits that confer parasite resistance (stabilized host-parasite-hyperparasite cycles). Under
241 conditions in which there is the influx of broad and complexed molecular parasite populations into
242 the habitat, the carrying capacity of the autocatalytic networks are exceeded, and a non-catalytic,
243 degenerative process of hyperparasite cycles ensues that leads to the appearance of antiparasitism,
244 and the birth of immunity in early life.

245 One of the strong features of this model is the solid empirical evidence supporting it. First of
246 all, the catalyst tradeoff analyses [38] applied to autocatalytic RNA networks predicts the emergence
247 of increasingly active ribozymes, with binding promiscuity and high catalytic activity being balanced
248 with specificity and relatively lower activity, which is precisely what the model shows. Prebiotic
249 models indeed support initially fast replicating communities with high molecular diversity, evolving
250 later into more faithful replicators with a narrower repertoire [42]. Moreover, tripartite microbial
251 populations composed of a host, a parasite and a hyperparasite are virtually ubiquitous in the
252 virosphere, as evidenced for bacteria-phages-phage satellites [24,25], host-RNA viruses-defective
253 interfering RNA viruses [26], and eukaryotic host-NCLDVs-virophages [27]. We previously developed a
254 hyperparasite framework based on the LV-equations that corroborates the superior homeostatic
255 viability of tripartite habitats [23]. It thus becomes likely, that early life network habitats with RNA
256 populations were constructed following this architectural logic.

257 Second, some sort of trans-hierarchical self-organization has to be stipulated at this stage of
258 the early RNA world [43], ideas conceptualized under various designations, such as prebiotic heredity
259 or ecology, and autopoiesis [35,36]. The early RNA world for instance implies lipid-vesicle mediated
260 compartmentalization and supramolecular selection [35]. Applied to our findings, ribozymes
261 conferring parasite resistance are selected early on, survive autopoietically within the quasispecies,
262 and can thus be reselected at later stages, once more efficient catalysts expose the community to
263 waves of threatening parasite invasions.

264 Third, it is striking that the appearance of microbial immunity is associated with
265 hyperparasites. Indeed, several recent studies show that phages and their satellites encode a set of
266 diverse antiphage systems [24,44], and defective interfering RNAs activate innate immunity [26,40].
267 Phage satellite-encoded immune systems protect bacteria from phage predation and were proposed
268 to be an integral innate immune component [45]. This notion could easily be extended to
269 omnipresent, bona fide hyperparasites, such as endogenous retroviruses [46,47]. In conclusion, our

model provides an overarching theoretical foundation for the notorious association of immunity with the parasite-hyperparasite sphere that is rooted in early life.

Furthermore, it should be noted that extant RNA viral quasispecies most closely resemble the primordial RNA populations described here. Large populations exhibiting fast replication recapitulate the fundamental tradeoff, favoring speed over polymerase accuracy [39]. High mutation rates enable the emergence and maintenance of substantial genetic variation, which leaves marks of the past ecological history within viral genomes via selective sweeps [48]. Moreover, in all contemporary tripartite models the host diverged to cellular life forms carrying polymerases phylogenetically distinct from those prevalent in viruses and selfish genetic elements [49], thus severely limiting genetic exchange. This represented a marked change from the primordial tripartite populations that freely swapped catalysts. The fact that the integration of the virosphere into early RNA world scenarios has been proposed as an important future agenda issue further valorizes our model [43].

Conceptually, our model capitalizes upon one recent and new school of thought in autocatalytic networks, namely the fact that chemical networks behave as ecological species interacting with one another [9]. Since the common underlying physical principle in catalysis and ecological interactions is the mass action law [8], it is not surprising that the biomathematical tools most accurately capturing the quantitative nature of these interactions are also equivalent, namely equations from the LV-family. Probably one of the most significant merits of our model founded in the equivalence of autocatalysis and ecological interactions is the fact that catalytic processes by their very nature can degenerate and become non-catalytic, and that a degenerative process has innovative power. We can summarize the sequence of events as, i) the catalytic process that is inherently error-prone and generates a quasispecies comprising molecular parasites; ii) sequential ecological interactions between host and parasite catalysts stabilize the habitat, forming tripartite ecological communities of host-parasites-hyperparasites; iii) a subgroup of hyperparasite catalysts in such habitats, when overwhelmed with invading parasite waves, can degenerate and create adaptive immunity against those parasites. Thus, in essence, degenerating catalysts create a new, homeostatic function that allows the habitat to adapt.

This model has by design intrinsic limitations, notably the fact that certain conclusions have yet to be experimentally verified [13,14]. Along the same line, it is impossible to infer causation from statistical associations, in particular the pervasive horizontal gene transfer associated with selfish genetic elements blurs the lines. These aspects will be addressed in future studies.

Methods or Models

Definitions and Terms

In what follows we assume the following: the Malthusian fitness r of a replicating (biochemical) system (or species) is defined by $\frac{d}{dt}R = \dot{R} = r\,R$, where R denotes the population size of the replication sys-tem or species. Evolutionary dynamics between two or more interacting species or systems, e.g. $R, S, P, \ldots$ are described by the following type of replicator differential equation e.g. for R:

$\dot{R}$ = "(A) terms of dynamics using habitat resources" x "(B) term of habitat resource restriction" + "(C) terms of dynamics not using habitat resources"

Terms A include both "terms of first order kinetics" and "terms of second order kinetics"

- "Terms of first order kinetics" are typically natural growth or decay rates, e.g. $\dot{R} = r\,R$. If $r > 0$, growth process described by rR eats into resources, whereas if $r < 0$, r is a decay rate, and de-cay dynamics does not.

  Terms of second order kinetics are derived from encounters between different species. For instance, the term *"bRS"* in *"$\dot{R} = -aR + bRS$"* has the following meaning: $bRS =$ $(probabiliy\ of\ an\ encounter\ between\ members\ of\ R\ and\ S)$ x $(impact\ of\ encounter\ on\ the\ fitness\ of\ R)$.

  Here we assume, according to the Lotka Volterra calculus that the probability of an encounter between members of R and S is proportional to the product RS.

  The term B of resource restriction is used in a habitat of limited resources. For instance, in the classic logistic equation $\dot{R} = r\,R\left(1 - \frac{R}{K}\right)$ the term $\left(1 - \frac{R}{K}\right) = B$ indicates that the habitat has limited resources supporting a maximum population K for R.

  Since every encounter involves resource consuming, we assume here, that all terms of second order kinetics are subject to resource restrictions and therefore need to be multiplied with the term B.

Example of a host-parasite interaction (R=host, P=parasite):

$$\dot{R} = (a\,R - bPR)\left(1 - \frac{R+P}{K}\right)$$

$$\dot{P} = -c\,P + dRP\left(1 - \frac{R+P}{K}\right)$$

where $a > 0, b > 0, c > 0, d > 0$. The first order dynamics growth process aR and all the second order processes eat into the habitat's resources, whereas the first order decay process $-c\,P$ does not. The habitat resources restriction is represented by the term $\left(1 - \frac{R+P}{K}\right)$, indicating that a maximum population size of $K = R + P$ can be nourished by the habitat and that both, S and P, compete for the same resources.

Definition of catalysts, ribozymes and enzymes

We use the term *catalyst*, sometimes also *perfect catalyst*, for any substance that accelerates the rate of a reaction without undergoing any change during the process. We call an *imperfect catalyst* a substance that accelerates the rate of a reaction, but that may undergo a change during the process; antibodies in the transition from hyperparasitism to antiparasitism are examples of imperfect catalysts (see process $(3')$, Suppl. Inf.). C*atalysts* displaying relatively higher substrate specificity generally have relatively lower catalytic activity, for instance the protoreplicase. A subgroup of catalysts is called here *enzymes*, if they are generated based on RNA templates. An enzyme may be either a *ribozyme*, if it is also used as an RNA template to generate other enzymes, or a *pure enzyme* if it cannot be used as an RNA template for other polymerization processes. The term *ribozyme* encompasses different enzymatic activities, such as recombinase, ligase and polymerase. Recombinases and ligases predominate in early protocycles, while polymerases appear only in advanced networks.

Model $\mathbf{1}^{(1)}$: habitat with one host autocatalytic cycle exposed to parasite invasion

$$\frac{dR_2}{dt} = a\ R_1 R_1 \Pi - d_{R_2} R_2$$
$$\frac{dR_1}{dt} = (b\ R_2 R_1 - \delta R_2 P)\Pi - d_{R_1} R_1$$
$$\frac{dP}{dt} = c R_2 P \Pi + \alpha - d_P P,$$

where a, b and c are positive constants, d_{R_i}, d_P are the decay rates of populations $R_i, i = 1, 2$, and of parasite P absent its host R_2; δ is parasite's P rate of exploitation of the host replicase R_2; and the habitat restriction term is defined as $\Pi = \left(1 - \frac{R_1 + R_2 + P/q_P}{K}\right)$, where $q_P > 1$, since typically parasites P are more efficient using the resources of the habitat than their hosts R. α is the parasite invasion parameter, i.e., per unit of time the habitat is invaded by α population units of parasites P.

Model $\mathbf{1}^{(0)}$habitat with primordial autocatalytic cycle, exposed to template invasion $\boldsymbol{P}$ (to become parasite later in evolution) and protoenzyme $\boldsymbol{E_1}$:

$$\frac{dP}{dt} = \alpha - d_P P$$
$$\frac{dE_1}{dt} = \beta - d_{E_1} E_1$$
$$\frac{dE_2}{dt} = a E_1 F_1 \Pi - d_{E_2} E_2$$
$$\frac{dF_1}{dt} = b E_2 P \Pi - d_{F_1} F_1,$$

where

E_2 is the protoreplicase produced by this primitive autocatalytic cycle,

F_1 is the prototemplate produced by this primitive autocatalytic cycle,

α is the inflow rate of RNA-template P into the habitat,

β is the inflow rate of protoenzyme/catalyst E_1 into the habitat,

a and b are positive constants,

d_{E_i}, d_{F_1}, d_P are the decay rates of enzyme populations $E_i, i = 1, 2$, of prototemplate d_{F_1}, and of RNA-template P,

$\Pi = \left(1 - \frac{E_2 + {}^{F_1}/_{q_{F_1}}}{K}\right)$ is the habitat restriction term where $q_{F_1} > 1$, since we assume that prototemplates F_1 are more efficient using the resources of the habitat than the more complex protoreplicase E_2.

Model $\mathbf{1}^{(\mathbf{1}+)}$: ***stabilized host-parasite-hyperparasite cycle***

$$\frac{dR_2}{dt} = (a\, R_1 R_1 - \delta_2 R_1 F_1)\Pi - d_{R_2} R_2$$
$$\frac{dR_1}{dt} = (b\, R_2 R_1 - \delta_1 R_2 P)\Pi - d_{R_1} R_1$$
$$\frac{dP}{dt} = (c R_2 P - \delta_3 F_2 P)\Pi + \alpha - d_P P$$
$$\frac{dF_1}{dt} = (d\, F_2 P + \beta R_2 P)\Pi - d_{F_1} F_1$$
$$\frac{dF_2}{dt} = (e\, R_1 F_1)\Pi - d_{F_2} F_2,$$

where

a, b, c, d, e are positive (reproduction) constants,

R_1 is a ribozyme of the autocatalytic host cycle,

R_2 is the replicase, i.e. another ribozyme (or may be also pure enzyme since it is not used as template) of the autocatalytic host cycle,

P is the parasite, exploiting replicase R_2 of the host cycle (exploitation parameter δ_1),

F_1 is the prototemplate exploiting the host R_1 (exploitation parameter δ_2) for protoreplicase F_2 production, triggered upon encounter of R_2 with P (trigger parameter β) and further produced by protoreplicase F_2 upon encounter with P,

F_2 is the protoreplicase (pure enzyme or ribozyme) producing F_1 upon encounter with parasite templates P ($F_2 = E_2$ from the primitive autocatalytic cycle of Model $1^{(0)}$), thus, exploiting template of parasite P (exploitation parameter δ_3),

d_{R_i}, d_{F_i}, d_P are the decay rates of populations $R_i, F_i, i = 1, 2$, and of P, absent new inflow or reproduction in the habitat,

α is the inflow rate of RNA template P into the habitat,

β is trigger parameter of replicase R_2 producing F_1 by encounter with parasite P,

$\Pi = \left(1 - \frac{R_1+R_2+{}^{P}/_{q_P}+{}^{F_1}/_{q_{F_1}}+{}^{F_2}/_{q_{F_2}}}{K}\right)$ is the habitat restriction term where $q_P, q_{F_1}, q_{F_2} > 1$, since we assume that parasitizing cycles of P and F_1, F_2 are more efficient using the resources of the habitat than the more complex host cycle R_1, R_2.

***Model* $\mathbf{1}^{(2)}$: integrated host-parasite-hyperparasite cycle without new parasite inflow**

The simulation results of Model $1^{(2)}$ are realized as follows: Model $1^{(1+)}$provides the starting basis, with arbitrary initial populations to establish a stabilized situation, given a moderate parasite inflow with $\alpha > 0$ (Fig. 3 lefthand part). Subsequently, Model $1^{(1+)}$ is adapted to Model $1^{(2)}$. All the parameters of Model $1^{(2)}$ are the same as those of Model $1^{(1+)}$, with the exception of the parasite inflow, now $\alpha = 0$. As initial conditions for the populations of Model $1^{(2)}$, we assume the terminal populations of the simulations of Model $1^{(2)}$, as calculated for Fig. 3 at the beginning of the righthand part. The larger fluctuations of the population at the beginning of the righthand part of Fig. 3 stem from the transition from Model $1^{(1+)}$ to Model $1^{(2)}$, where the parasite inflow abruptly decreases from $\alpha > 0$ to $\alpha = 0$. The fact that parasites in Model $1^{(2)}$ are now reproduced without new inflow shows that they can be viewed as fully integrated into the autocatalytic network, realized in $1^{(2)}$ (see Suppl. Inf.).

Model* $\mathbf{1}^{(1-)}$: *transition from hyperparasitism to antiparasitism

$$\frac{dR_2}{dt} = (a\ R_1R_1 - \delta_2R_1F_1)\Pi - d_{R_2}R_2$$
$$\frac{dR_1}{dt} = (b\ R_2R_1 - \delta_1R_2P)\Pi - d_{R_1}R_1$$
$$\frac{dP}{dt} = (cR_2P - \delta_3F_2P)\Pi + \alpha - \theta_PF_2P - d_PP$$
$$\frac{dF_1}{dt} = (d\ F_2P + \beta R_2P)\Pi - d_{F_1}F_1$$
$$\frac{dF_2}{dt} = (e\ R_1F_1)\Pi - \theta_{F_2}F_2P - d_{F_2}F_2,$$

Where all the parameters are the same as in Model $1^{(1+)}$, with the exception that now two additional parameters are added, namely θ_P and θ_{F_2}. The term θ_PF_2P denotes the number of population units of parasites P made unviable per unit of time during the encounter of protoreplicase F_2 with P, now also functioning as antibody. In addition, the term $\theta_{F_2}F_2P$ denotes the number of population units of protoreplicase/antibody F_2 made unviable per unit of time during the encounter of

409 protoreplicase/antibody F_2 with P. For instance, if $\theta_p = \theta_{F_2}$, then the same number of antibodies and
410 parasites are made unviable during their encounter. In the evolutionary transition from
411 hyperparasitism to antiparasitism F_2 can still act both ways, as imperfect protoreplicase and as
412 antibody directed against the parasite P. We call it *imperfect* protoreplicase or imperfect catalyst,
413 because it still produces/catalyzes prototemplate F_1 upon encounter with template P. But now F_2
414 may be damaged during this process, therefore cannot be called a *perfect* catalyst anymore. The
415 replication differential equations of Model $1^{(1-)}$ realizing a pure antiparasitism situation are the
416 same as in Model $1^{(1-)}$ of Fig. 4A, except that the reproduction parameters hyperparasite d and δ_3
417 are now set to zero: $\delta_3 = d = 0$.
418
419
420 **Acknowledgements**
421 We thank Thomas Curran for discussions.
422
423
424 **References**
425 1. Betts HC, Puttick MN, Clark JW, Williams TA, Donoghue PCJ, Pisani D. Integrated genomic and
426 fossil evidence illuminates life's early evolution and eukaryote origin. Nat Ecol Evol 2018; 2:
427 1556-1562. doi: 10.1038/s41559-018-0644-x. Epub 2018 Aug 20. PMID: 30127539
428 2. Jeancolas C, Malaterre C, Nghe P. Thresholds in Origin of Life Scenarios. iScience 2020; 23:
429 101756. doi: 10.1016/j.isci.2020.101756. PMID: 33241201
430 3. Kauffman SA. Autocatalytic sets of proteins. J Theor Biol 1986; 119: 1-24. doi: 10.1016/s0022-
431 5193(86)80047-9. PMID: 3713221
432 4. Hordijk W, Hein J, Steel M. Autocatalytic Sets and the Origin of Life. Entropy. 2010; 12: 1733-
433 1742. https://doi.org/10.3390/e12071733
434 5. Shapiro R. Small molecule interactions were central to the origin of life. Q Rev Biol. 2006; 81:
435 105-25. doi: 10.1086/506024. PMID: 16776061
436 6. Ruiz-Mirazo K, Briones C, de la Escosura A. Prebiotic systems chemistry: new perspectives for the
437 origins of life. Chem Rev 2014: 114: 285-366. doi: 10.1021/cr2004844. Epub 2013 Oct 31. PMID:
438 24171674
439 7. Peng Z, Linderoth J, Baum DA. The hierarchical organization of autocatalytic reaction networks
440 and its relevance to the origin of life. PLoS Comput Biol 2022; 18: e1010498. doi:
441 10.1371/journal.pcbi.1010498. PMID: 36084149

442 8. Voit EO, Martens HA, Omholt SW. 150 years of the mass action law. PLoS Comput Biol 2015;
443 11:e1004012. doi: 10.1371/journal.pcbi.1004012. PMID: 25569257
444 9. Peng Z, Plum AM, Gagrani P, Baum DA. An ecological framework for the analysis of prebiotic
445 chemical reaction networks. J Theor Biol 2020; 507: 110451. doi: 10.1016/j.jtbi.2020.110451.
446 Epub 2020 Aug 12. PMID: 32800733
447 10. Joyce GF, Szostak JW. Protocells and RNA Self-Replication. Cold Spring Harb Perspect Biol 2018;
448 10: a034801. doi: 10.1101/cshperspect.a034801. PMID: 30181195
449 11. Kauffman SA. Cellular homeostasis, epigenesis and replication in randomly aggregated
450 macromolecular systems. J Cybernet 1971; 1: 71-96.
451 12. Eigen M. Self-organization of matter and the evolution of biological macromolecules.
452 Naturwissenschaften 1971; 58: 465-523.
453 13. Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N. Spontaneous network
454 formation among cooperative RNA replicators. Nature 2012; 491: 72-7. doi:
455 10.1038/nature11549. Epub 2012 Oct 17. PMID: 23075853
456 14. Tjhung KF, Shokhirev MN, Horning DP, Joyce GF. An RNA polymerase ribozyme that synthesizes
457 its own ancestor. Proc Natl Acad Sci U S A 2020; 117: 2906-2913. doi: 10.1073/pnas.1914282117.
458 Epub 2020 Jan 27. PMID: 31988127
459 15. Pavlinova P, Lambert CN, Malaterre C, Nghe P. Abiogenesis through gradual evolution of
460 autocatalysis into template-based replication. FEBS Lett 2023; 597: 344-379. doi: 10.1002/1873-
461 3468.14507. Epub 2022 Nov 17. PMID: 36203246
462 16. Xavier JC, Hordijk W, Kauffman S, Steel M, Martin WF. Autocatalytic chemical networks at the
463 origin of metabolism. Proc Biol Sci 2020; 287: 20192377. doi: 10.1098/rspb.2019.2377. Epub
464 2020 Mar 11. PMID: 32156207
465 17. Stubbs RT, Yadav M, Krishnamurthy R, Springsteen G. A plausible metal-free ancestral analogue
466 of the Krebs cycle composed entirely of α-ketoacids. Nat Chem 2020; 12: 1016-1022. doi:
467 10.1038/s41557-020-00560-7. Epub 2020 Oct 12. PMID: 33046840
468 18. Takeuchi N, Hogeweg P, Koonin EV. On the origin of DNA genomes: evolution of the division of
469 labor between template and catalyst in model replicator systems. PLoS Comput Biol 2011;
470 7:e1002024. doi: 10.1371/journal.pcbi.1002024. Epub 2011 Mar 24. PMID: 21455287
471 19. Domingo E, García-Crespo C, Perales C. Historical Perspective on the Discovery of the
472 Quasispecies Concept. Annu Rev Virol 2021; 8: 51-72. doi: 10.1146/annurev-virology-091919-
473 105900. PMID: 34586874
474 20. Bull JJ, Meyers LA, Lachmann M. Quasispecies made simple. PLoS Comput Biol 2005; 1: e61. doi:
475 10.1371/journal.pcbi.0010061. PMID: 16322763

476 21. Koonin EV, Wolf YI, Katsnelson MI. Inevitability of the emergence and persistence of genetic
477 parasites caused by evolutionary instability of parasite-free states. Biol Direct 2017; 12: 31. doi:
478 10.1186/s13062-017-0202-5. PMID: 29202832
479 22. Iranzo J, Puigbò P, Lobkovsky AE, Wolf YI, Koonin EV. Inevitability of Genetic Parasites. Genome
480 Biol Evol 2016; 8: 2856-2869. doi: 10.1093/gbe/evw193. PMID
481 23. Conrad B, Iseli C, Pirovino M. Energy-harnessing problem solving of primordial life: Modeling the
482 emergence of catalytic host-nested parasite life cycles. PLoS One 2023; 18: e0281661. doi:
483 10.1371/journal.pone.0281661. PMID: 36972235
484 24. de Sousa JAM, Fillol-Salom A, Penadés JR, Rocha EPC. Identification and characterization of
485 thousands of bacteriophage satellites across bacteria. Nucleic Acids Res 2023; 51: 2759-2777.
486 doi: 10.1093/nar/gkad123. PMID: 36869669
487 25. Eppley JM, Biller SJ, Luo E, Burger A, DeLong EF. Marine viral particles reveal an expansive
488 repertoire of phage-parasitizing mobile elements. Proc Natl Acad Sci U S A 2022; 119:
489 e2212722119. doi: 10.1073/pnas.2212722119. Epub 2022 Oct 18. PMID: 36256808
490 26. Vignuzzi M, López CB. Defective viral genomes are key drivers of the virus-host interaction. Nat
491 Microbiol 2019; 4: 1075-1087. doi: 10.1038/s41564-019-0465-y. Epub 2019 Jun 3. PMID:
492 31160826
493 27. Katzourakis A, Aswad A. The origins of giant viruses, virophages and their relatives in host
494 genomes. BMC Biol 2014; 12: 51. doi: 10.1186/s12915-014-0051-y. PMID: 25184667
495 28. Davis JD, Olivença DV, Brown SP and Voit EO. Methods of quantifying interactions among
496 populations using Lotka-Volterra models. Front Syst Biol 2022; 2: 1021897. doi:
497 10.3389/fsysb.2022.1021897
498 29. Lotka AJ. Contribution to the Theory of Periodic Reactions. J Phys Chem 1910; 14: 271–274.
499 30. Boros B, Hofbauer J, Müller S, Regensburger G. The Center Problem for the Lotka Reactions with
500 Generalized Mass-Action Kinetics. Qual Theory Dyn Syst 2018; 17: 403-410. doi:
501 10.1007/s12346-017-0243-2. Epub 2017 Jun 13. PMID: 30636938
502 31. Michaelis L, Menten ML, Johnson KA, Goody RS. The original Michaelis constant: translation of
503 the 1913 Michaelis-Menten paper. Biochemistry 2011; 50: 8264-9. doi: 10.1021/bi201284u.
504 Epub 2011 Sep 9. PMID: 21888353
505 32. Kooi BW, Poggiale JC. Modelling, singular perturbation and bifurcation analyses of bitrophic food
506 chains. Math Biosci 2018; 301: 93-110. doi: 10.1016/j.mbs.2018.04.006. Epub 2018 Apr 22.
507 PMID: 29684407
508 33. Poggiale JC, Aldebert C, Girardot B, Kooi BW. Analysis of a predator-prey model with specific time
509 scales: a geometrical approach proving the occurrence of canard solutions. J Math Biol 2020; 80:
510 39-60. doi: 10.1007/s00285-019-01337-4. Epub 2019 Feb 20. PMID: 30788562

511 34. Eilertsen J, Roussel MR, Schnell S, Walcher S. On the quasi-steady-state approximation in an
512 open Michaelis-Menten reaction mechanism. AIMS Math 2021; 6: 6781-6814. doi:
513 10.3934/math.2021398. Epub 2021 Apr 21. PMID: 34142000
514 35. Wieczorek R. On prebiotic ecology, supramolecular selection and autopoiesis. Orig Life Evol
515 Biosph 2012; 42:445-50; discussion 451-2. doi: 10.1007/s11084-012-9306-1. PMID: 23132761
516 36. Igamberdiev AU, Kleczkowski LA. Toward understanding the emergence of life: A dual function of
517 the system of nucleotides in the metabolically closed autopoietic organization. Biosystems 2023;
518 224: 104837. doi: 10.1016/j.biosystems.2023.104837. Epub 2023 Jan 14. PMID: 36649884
519 37. Janzen E, Blanco C, Peng H, Kenchel J, Chen IA. Promiscuous Ribozymes and Their Proposed Role
520 in Prebiotic Evolution. Chem Rev 2020; 120: 4879-4897. doi: 10.1021/acs.chemrev.9b00620.
521 Epub 2020 Feb 3. PMID: 32011135
522 38. Tawfik DS. Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic
523 efficiency? Curr Opin Chem Biol 2014; 21:73-80. doi: 10.1016/j.cbpa.2014.05.008. Epub 2014 Jun
524 20. PMID: 24954689
525 39. Fitzsimmons WJ, Woods RJ, McCrone JT, Woodman A, Arnold JJ, Yennawar M, Evans R, Cameron
526 CE, Lauring AS. A speed-fidelity trade-off determines the mutation rate and virulence of an RNA
527 virus. PLoS Biol 2018; 16: e2006459. doi: 10.1371/journal.pbio.2006459. PMID: 29953453
528 40. Genoyer E, López CB. The Impact of Defective Viruses on Infection and Immunity. Annu Rev Virol.
529 2019; 6: 547-566. doi: 10.1146/annurev-virology-092818-015652. Epub 2019 May 13. PMID:
530 31082310).
531 41. Ibarra-Chávez R, Brady A, Chen J, Penadés JR, Haag AF. Phage-inducible chromosomal islands
532 promote genetic variability by blocking phage reproduction and protecting transductants from
533 phage lysis. PLoS Genet 2022; 18: e1010146. doi: 10.1371/journal.pgen.1010146. PMID:
534 35344558
535 42. Markovitch O, Lancet D. Multispecies population dynamics of prebiotic compositional
536 assemblies. J Theor Biol 2014; 357: 26-34. doi: 10.1016/j.jtbi.2014.05.005. Epub 2014 May 14.
537 PMID: 24831416
538 43. Fine JL, Pearlman RE. On the origin of life: an RNA-focused synthesis and narrative. RNA 2023;
539 29: 1085-1098. doi: 10.1261/rna.079598.123. Epub 2023 May 4. PMID: 37142437
540 44. Rousset F, Depardieu F, Miele S, Dowding J, Laval AL, Lieberman E, Garry D, Rocha EPC, Bernheim
541 A, Bikard D. Phages and their satellites encode hotspots of antiviral systems. Cell Host Microbe
542 2022; 30: 740-753.e5. doi: 10.1016/j.chom.2022.02.018. Epub 2022 Mar 21. PMID: 35316646
543 45. Fillol-Salom A, Miguel-Romero L, Marina A, Chen J, Penadés JR. Beyond the CRISPR-Cas
544 safeguard: PICI-encoded innate immune systems protect bacteria from bacteriophage predation.

Curr Opin Microbiol 2020; 56: 52-58. doi: 10.1016/j.mib.2020.06.002. Epub 2020 Jul 9. PMID: 32653777

46. Stauffer Y, Marguerat S, Meylan F, Ucla C, Sutkowski N, Huber B, Pelet T, Conrad B. Interferon-alpha-induced endogenous superantigen. a model linking environment and autoimmunity. Immunity 2001; 15: 591-601. doi: 10.1016/s1074-7613(01)00212-6. PMID: 11672541
47. Malfavon-Borja R, Feschotte C. Fighting fire with fire: endogenous retrovirus envelopes as restriction factors. J Virol 2015; 89: 4047-50. doi: 10.1128/JVI.03653-14. Epub 2015 Feb 4. PMID: 25653437
48. Zhao L, Seth-Pasricha M, Stemate D, Crespo-Bellido A, Gagnon J, Draghi J, Duffy S. Existing Host Range Mutations Constrain Further Emergence of RNA Viruses. J Virol 2019; 93: e01385-18. doi: 10.1128/JVI.01385-18. PMID: 30463962
49. Koonin EV, Dolja VV. Virus world as an evolutionary network of viruses and capsidless selfish elements. Microbiol Mol Biol Rev 2014; 78: 278-303. doi: 10.1128/MMBR.00049-13. PMID: 24847023

Supporting information captions

Self-replicating RNA molecules

Let us consider a set of populations $\{R_1, R_2, \dots, R_n\}$ of different RNA molecules or enzymes in a primordial RNA setting. Enzymatic functions are exerted by RNA molecules called ribozymes. Ribozymes and other enzymes are forming a catalytic cycle, if

$$R_1 \rightarrow^c R_2 \rightarrow^c R_3 \rightarrow^c \dots \rightarrow^c R_n \rightarrow^c R_1,$$

where the symbol "$A \rightarrow^c B$" means: A catalyzes B.

In this setting, RNA A can display two different functions, the coding and the enzymatic function. RNA-molecules without enzymatic activity have at least a coding potential. Let us call the coding function "template" of RNA molecule. A pure enzyme not bearing a coding function lacks the ability to serve as template. Let us further assume that in this instance successful enzymatic activities require both enzyme and template. An enzymatic activity is thus exerted by a ribozyme or a pure enzyme upon encounter with an RNA template. For example, consider the following equations:

$$\frac{dR_2}{dt} = a\ R_1 R_1 \Pi - d_{R_2} R_2 \qquad (1.1)$$

$$\frac{dR_1}{dt} = b\ R_2 R_1 \Pi - d_{R_1} R_1 \qquad (1.2),$$

where a and b are positive constants, d_{R_i} are the decay rates of populations R_i, $i = 1, 2$, and

$$\Pi = \left(1 - \frac{R_1 + R_2}{K}\right)$$

the habitat restriction constant. K is the maximum number of R_1 or R_2 molecules, the habitat can feed.

Replication equations (1.1), (1.2) form a two-step autocatalytic cycle. In (1.1), R_1 serves both as enzyme and template, in (1.2) R_2 serves as enzyme and R_1 as template, i.e. upon encounter of two different molecules R_1, one serving as template and the other as enzyme, a ribozyme or pure enzyme R_2 is produced/catalyzed, and upon encounter of a molecule R_1 serving as template with a molecule R_2 serving as enzyme the original ribozyme R_1 is produced, forming a complete replication cycle. We symbolize this autocatalytic process by:

$$R_1 \Rightarrow (R_1 \leftrightarrow R_1) \rightarrow^c R_2 \Rightarrow (R_2 \leftrightarrow R_1) \rightarrow^c R_1 \quad (1),$$

where the symbol "⇒" means "*enables*" and the symbol "$A \leftrightarrow B$" means "*encounter of A with B*" and, as already stated above the symbol "$\rightarrow^c$" means "*catalyzes*".

Thus, autocatalytic cycles can take the form:

$$R_1 \Rightarrow (R_1 \leftrightarrow R_1) \rightarrow^c R_2 \Rightarrow (R_2 \leftrightarrow R_1) \rightarrow^c R_3 \Rightarrow (R_3 \leftrightarrow R_1) \rightarrow^c \ldots \rightarrow^c R_n \Rightarrow (R_n \leftrightarrow R_1) \rightarrow^c R_1.$$

In this example, only R_1 is used as a template during every enzymatic subactivity of the autocatalytic reproduction cycle; every enzymatic activity produces a new ribozyme or a new pure enzyme $R_i, i = 1, .. n$.

Rather than producing a further ribozyme or a pure enzyme, an enzymatic activity can also produce a further template, e.g.

$$R_1 \Rightarrow (R_1 \leftrightarrow R_1) \rightarrow^c R_2 \Rightarrow (R_2 \leftrightarrow R_1) \rightarrow^c R_3 \Rightarrow (R_2 \leftrightarrow R_3) \rightarrow^c R_4 \Rightarrow (R_4 \leftrightarrow R_1) \rightarrow^c R_1$$

reads as follows: a ribozyme R_1 enables the encounter with another copy of itself, one serving as enzyme, the other as template, and both producing a new enzyme R_2. Upon encounter with a template R_1, R_2 catalyzes a new template R_3, and upon encounter with this new template a new enzyme R_4, which in turn catalyzes the final reproduction of a new ribozyme molecule R_1 using another molecule R_1 as template.

Parasitism in autocatalytic cycles

To outline the problem of parasitism, e.g. in the autocatalytic cycle (1), the ribozyme R_2, which may also be a pure enzyme, is acting as a replicase for R_1. Such a replicase is promiscuous and highly active, therefore, R_2 will inevitably be parasitized by parasitic RNA-templates P:

$$\frac{dP}{dt} = cR_2P\Pi - d_P P \qquad (2.1),$$

or, written in above introduced technology template format:

$$P \Rightarrow (R_2 \leftrightarrow P) \rightarrow^c P \quad (2).$$

Here again, c is a positive constant, $d_P > 0$ is the decay rate of parasite P absent its host R_2, and the habitat restriction term Π extends to

$$\Pi = \left(1 - \frac{R_1 + R_2 + {}^{P}/_{q_P}}{K}\right),$$

604 where $q_P > 1$, since typically parasites P are more efficient at using the resources of the habitat
605 compared to their hosts R. A ribozyme or pure enzyme molecule R_2 that is used as the replicase for P
606 cannot at the same time be used as the replicase for R_1, thus P parasitizes the whole autocatalytic
607 cycle (1).
608
609 ***A stable, primordial autocatalytic cycle*** (1^{origin}) ***in the presence of parasites***
610 Let us assume the primordial habitat has an abundant inflow of RNA templates P of relatively
611 advanced complexity from external sources, and also an abundant inflow of a primary protoenzyme
612 E_1. In this mixture, containing P and E_1, we assume that the following primordial autocatalytic cycle
613 has already evolved. The cycle produces a prototemplate F_1 and a protoreplicase E_2 obeying the
614 following autoca-talytic equations:

$$\frac{dE_2}{dt} = e\ E_1 F_1 \Pi - d_{E_2} E_2 \qquad (1.1^{origin})$$

$$\frac{dF_1}{dt} = f\ E_2 P\ \Pi - d_{F_1} F_1 \qquad (1.2^{origin}),$$

615 where e and f are positive constants and d_{E_2}, d_{F_1} are, again decay rates as the ones above. And let
616 the habitat restriction term Π be accordingly defined to fit this situation. Thus, protoenzyme E_1 and
617 temp-lates P are given through inflow into the habitat, and ribozyme E_2, which may also be a pure
618 enzyme bearing no template function, and prototemplate F_1 are autocatalytically reproduced.
619 This original primitive autocatalytic process reads in above format as follows:

$$F_1 \Rightarrow (E_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow (E_2 \leftrightarrow P) \rightarrow^c F_1 \qquad (1^{origin}),$$

620 The protoenzyme E_1 produces a protoreplicase E_2 encoded by prototemplate F_1. The fact that E_2 is
621 only a protoreplicase means that is not yet capable of efficiently replicating a complex RNA template
622 P. But provided P is above a given threshold complexity, the protoreplicase E_2 can at least produce
623 copies of the prototemplate F_1. As long as the habitat has a minimal threshold concentration of both,
624 acceptably complexed templates P and protoenzymes E_1, the archaic autocatalytic cycle (1^{origin})
625 can thrive and produce further copies of F_1 and E_2. Remarkably, by design the primitive autocatalytic
626 cycle is immune against parasite invasion by P trying to exploit the protoreplicase E_2 for its own
627 replication. On the contrary, the primordial autocatalytic cycle (1^{origin}) depends on template
628 invasion P – in other words parasite invasion – for its own reproduction.
629
630 ***Evolution of the primordial autocatalytic cycle*** $(\mathbf{1}^{origin})$ ***into a full-fledged autocatalytic cycle*** (1)

We surmise that the host cycle (1) is capable of prebiotic heredity [10-12], that the primordial autoca-talytic cycle (1^{origin}) has been conserved in the quasispecies, and therefore can again be selected for upon parasite invasion. Therefore, the protoreplicase E_2 can be reselected from the quasispecies. This triggered process can take the following form:

$$R_2 \Rightarrow (R_2 \leftrightarrow P) \rightarrow^c F_1 \Rightarrow (R_1 \leftrightarrow F_1) \rightarrow^c E_2 \quad (1\&2 \rightarrow 3)$$

The fully functional replicase R_2 is designed to replicate R_1 according to (1.2), but it can also be exploited to replicate genetic parasites P according to (2). Upon encounter with a parasitic template P, it not only replicates P but also produces via its catalyst subunit E_2 a prototemplate F_1, as did its parent E_2 replicase. What is true for the replicase R_2 – that it is more efficient than its parent catalyst, the protoreplicase E_2 – is also true for ribozyme R_1. It is catalytically more efficient than protoenzyme E_1 it had replaced during evolution. R_1 can not only produce a full-fledged replicase R_2 given another copy of itself as a template, but via its protoenzyme E_1-subunit it reproduces the protoreplicase E_2 upon encounter with prototemplate F_1:

$$F_1 \Rightarrow (R_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow (E_2 \leftrightarrow P) \rightarrow^c F_1 \quad (3).$$

In order to name the processes:

(1) is the host cycle

(2) is the parasite cycle exploiting the host cycle

(3) is the hyperparasite cycle, parasitizing the host (exploiting ribozyme R_1 for protoreplicase E_2 production), and parasitizing the parasite P (exploiting its template for prototemplate F_1 production) triggered throughout the process $(1\&2 \rightarrow 3)$.

Prebiotic heredity of autocatalytic cycles to historical parasite exposure

In an environment with a constant – i.e. moderate, thus manageable – inflow of invasive parasites the full-fledged autocatalytic host-parasite-hyperparasite cycle $(1),(2),(3)$ together with the trigger process $(1\&2 \rightarrow 3)$ becomes increasingly integrated. If the nature of the constant inflow is always of the same special type of invasive parasites $P^{special}$, then the hyperparasite process (3) as well as the trigger process $(1\&2 \rightarrow 3)$ can mutate/adapt to processes that are fully adapted to $P^{special}$:

$$\left(3^{special}\right): \quad F_1^{special} \Rightarrow \left(R_1 \leftrightarrow F_1^{special}\right) \rightarrow^c E_2^{special} \Rightarrow \left(E_2^{special} \leftrightarrow P^{special}\right) \rightarrow^c F_1^{special}$$

$$\left(1\&2 \rightarrow 3^{special}\right): \; R_2 \Rightarrow \left(R_2 \leftrightarrow P^{special}\right) \rightarrow^c F_1^{special} \Rightarrow \left(R_1 \leftrightarrow F_1^{special}\right) \rightarrow^c E_2^{special}$$

It is reasonable to assume that the host-parasite-hyperparasite process $(1),(2),(3)$ together with the trigger process $(1\&2 \rightarrow 3)$ mutates under permanent special parasite inflow $P^{special}$ to an adapted process $(1),(2),(3^{special})$, together with the adapted trigger process $\left(1\&2 \rightarrow 3^{special}\right)$. This network of newly formed autocatalytic cycles reproduces all its elements. All adapted elements

660 $R_1, R_2, P^{special}, F_1^{special}, E_2^{special}$ reproduce, even when the special external parasites $P^{special}$ cease
661 to invade the habitat.
662 With the above we now can say: the primordial autocatalytic cycle $(1)^{(0)}$ has mutated to the
663 cycle $(1)^{(1)}$, which in turn mutates to the cycle $(1)^{(2)}$, after having faced a longer term invasion of
664 the same type of parasites $P^{special}$.
665 With this setup evolution exhibits two salient properties:
666 - The systems $(1)^{(0)}$, $(1)^{(1)}$, and $(1)^{(2)}$ can be viewed as increasing complexed networks of
667 autonomous autocatalytic cycles. They can be seen as the first steps in kickstarting evolution
668 towards complexed living systems.
669 - Prebiotic heredity [10-12] allows the cycle-network to preserve the complete record of its
670 historical environmental exposure to parasites:
671 o $(1)^{(2)}$ preserves the record of $(1)^{(1)}$, especially the type of its infection $P^{special}$, and
672 o $(1)^{(1)}$ preserves the record of its origin $(1)^{(0)}$ being immune against a broad range of
673 para-site infections P, and
674 o $(1)^{(0)}$ preserves the record of immunity against a broad range of parasite infections P.
675 As evolution progresses, the autocatalytic cycles $(1)^{(2)}$ and $(1)^{(1)}$ can now be invaded by new
676 parasites, which can be integrated into more complexed systems, e.g. $(1)^{(3)}$ or $(1)^{(2')}$, which in turn
677 always keep track of the entire historical parasite exposure.
678
679 ***Emergence of antiparasitism in autocatalytic cycles***
680 Let us assume the autocatalytic cycle $(1)^{(1)}$ is permanently invaded by many different types of
681 parasites P, to which it was not yet been exposed historically. The integration of every special
682 subtype of these parasites would not be viable and lead to a complexity overflow in the evolving
683 autocatalytic network. As first line of defense, the emergence of the autocatalytic host-parasite-
684 hyperparasite cycle $(1)^{(1+)} = \{(1), (2), (3), (1\&2 \rightarrow 3)\}$, as previously described would be able to
685 stabilize the situation, if infection P is not too overwhelming, and as time advances not too diverse.
686 For a more sustained, overwhelming and diverse parasite attack, another type of immunity would be
687 adapted, although it also originates from the already established host-parasite-hyperparasite
688 stabilization $(1)^{(1+)}$:
689 In the original primitive autocatalytic cycle (1^{origin}), $F_1 \Rightarrow (E_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow (E_2 \leftrightarrow$
690 $P) \rightarrow^c F_1$, it is essential the both encounters $(E_1 \leftrightarrow F_1)$ and $(E_2 \leftrightarrow P)$ preserve their catalytic nature
691 during the whole process, since otherwise – e.g. if they are damaged, this will heavily reduce their
692 respective population sizes – the autocatalytic nature of (1^{origin}) could break down rapidly.
693 However, this is not the case for a full-fledged autocatalytic host-parasite-hyperparasite trio

694 $(1),(2),(3)$ in $(1)^{(1+)}$ that has already evolved. The hyperparasite $(3), F_1 \Rightarrow (R_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow$
695 $(E_2 \leftrightarrow P) \rightarrow^c F_1$, represents the host's (1) evolutionary heredity from its primordial origin (1^{origin}).
696 If in this hyperparasite cycle (3), the encounter of e.g. the protoenzyme E_2 with the genetic parasite
697 P loses its catalytic nature, then the autocatalytic nature of the full-fledged cycle $(1),(2),(3)$ in
698 $(1)^{(1+)}$ would not break down. If P, or both P and E_2 are damaged during their encounter, this would
699 only lead to a strengthening of the host cycle (1), which is viable also in the absence of parasite and
700 hyperparasite cycles $(2),(3)$. As a result, the cycles $(2),(3)$ would not break down, given the fact
701 that the habitat still faces a sustained genetic parasite invasion P. Thus, the full-fledged cycle
702 $(1),(2),(3)$ would easily survive if the encounter of the protoreplicase E_2 with the genetic parasite P
703 would lose its catalytic nature. On the contrary, if P would be damaged, the stability, i.e. the fitness of
704 $(1)^{(1)}$ and with it $(1)^{(1+)}$, would increase further. Under such circumstances, we expect evolution to
705 establish a degenerative process of the $(E_2 \leftrightarrow P)$ encounter. We write for such a degenerated
706 encounter:

$$E_2 \leftrightarrow^d P,$$

707 meaning that during the encounter of E_2 with P, both elements, E_2 and P, or at least the
708 template P, are damaged and made unviable:

$$F_1 \Rightarrow (R_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow \left(E_2 \leftrightarrow^d P\right) \rightarrow F_1 \quad (3')$$

709 Here the symbol "→" means "produces". But this "production" is not catalytic anymore, since the
710 "producers" are damaged during the process. For an illustration of this situation where hyper-
711 parasitism coexists with antiparasitism (i.e. F_1 production upon encounter of E_2 with P exploiting
712 P as template) see Fig. 4A and Fig. 5.

713 The antiparasite process can now degenerate even further to $(3'')$, where the encounter
714 of prototemplate E_2, which we now call *antibodies,* with parasite P loses its ability to produce
715 the prototemplate F_1:

$$F_1 \Rightarrow (R_1 \leftrightarrow F_1) \rightarrow^c E_2 \Rightarrow \left(E_2 \leftrightarrow^d P\right) \quad (3'')$$

716 The ability of the antibody E_2 to produce a prototemplate F_1 upon encounter with the parasite
717 template P is not an evolutionary necessary one in this situation. Prototemplates F_1 are only
718 needed as long as parasites P exploit replicase R_2 for reproduction, but replicase R_2 already
719 triggers F_1-production upon encounter with parasites P through the trigger process $(1\&2 \rightarrow 3)$.
720 Therefore, F_1-production through the antiparasite process $(3'')$ becomes redundant.
721
722
723 **Author contributions**

Conception and design of the work: MP and BC. Data acquisition MP; data analysis MP, CI, BC; interpretation of data MP, CI, JAC, BC. Creation of new software used in the work: MP, CI. Drafted the work and revised it: BC, MP and JAC.

Data availability

The data, including scripts, is fully made available at the following sites (https://github.com/BICC-UNIL-EPFL/parasite-immunity-from-hyperparasitism [github.com]; https://doi.org/10.5281/zenodo.7970788 [doi.org]).

Competing interests

The authors declare no competing interests.

Figures and Figure Legends

Fig. 1

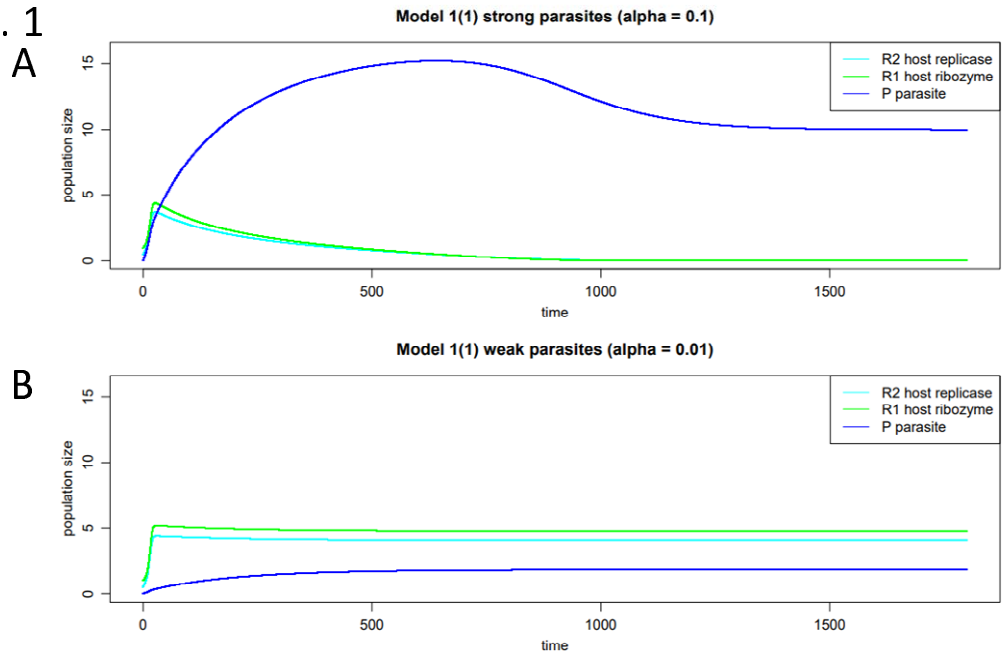

Figure 1. Habitat with one autocatalytic host cycle, variably parasite exposed

Fig. 1A. The catalysts in an autocatalytic cycle (ribozyme R_2, template R_1) exposed to incoming parasites P are promiscuous and hence also amplify P. Under high parasite exposure, parasites will therefore outcompete host molecules.

Fig. 1B. Under weak parasite exposure, the host molecules R_1 and R_2 prevail.

Fig. 2

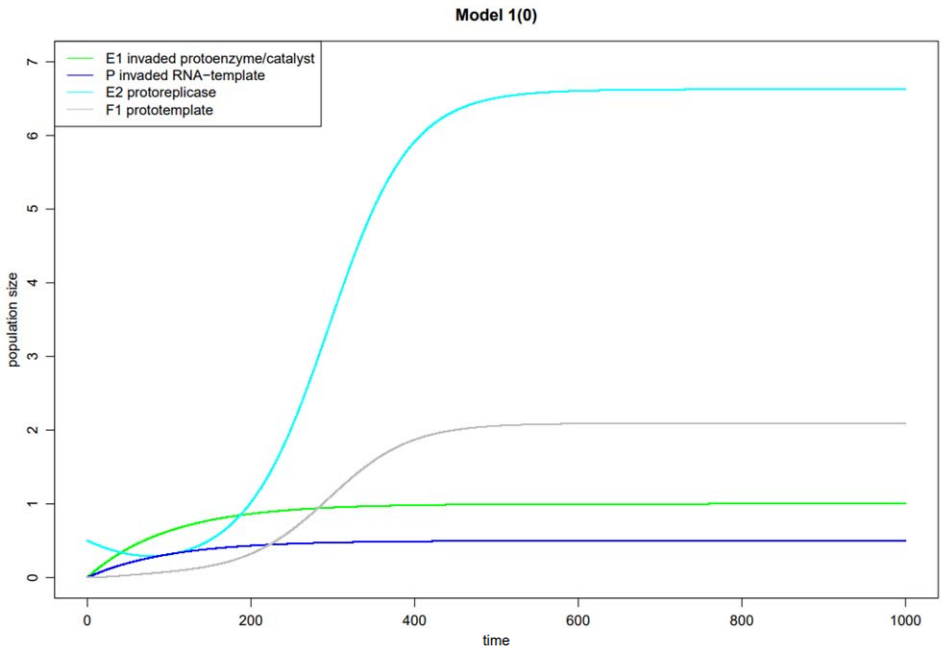

Figure 2. Habitat with primordial autocatalytic cycle

Parasite-specific catalysts with relatively lower catalytic activity in the habitat (E_2 protoreplicase) will confer parasite resistance and thus ensure abundance of habitat molecules in the presence of parasite inflow. Host molecules are expectedly less prevalent compared to Fig. 1B, at least prototemplate F_1.

Fig. 3

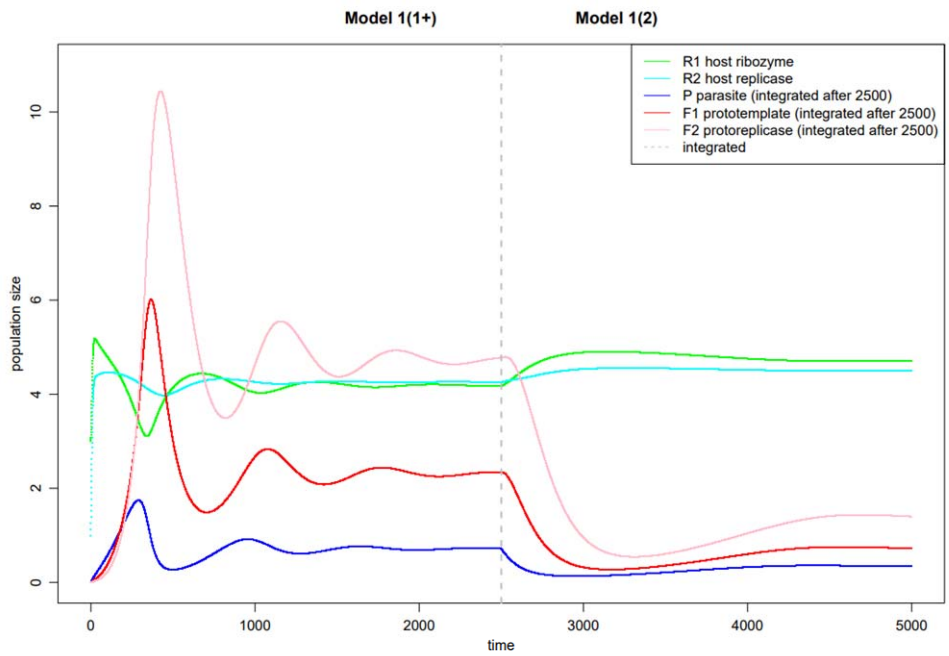

Figure 3. Stabilized host-parasite-hyperparasite cycle with variable parasite inflow

Fig. 3 left panel (Model 1(1+)). A habitat with highly efficient catalysts (self-replicase R_2, template R_1) becomes stabilized with moderate parasite inflow via its relatively less active catalyst-subunit E_2 that confers parasite resistance. F_2 (formerly E_2 in Fig. 2) is generated as well, it acts additively with *R2* and will be important for antiparasitism.

Fig. 3 right panel (Model 1(2)). Under constant, moderate and thus manageable parasite inflow ($P^{special}$, Suppl. Inf.), the hyperparasite cycle mutates and becomes integrated into the habitat, even when parasite inflow ceases (prebiotic heredity and mutation-selection balance). Host molecule prevalence is comparable to a habitat without parasites (Fig. 1B).

Fig. 4A

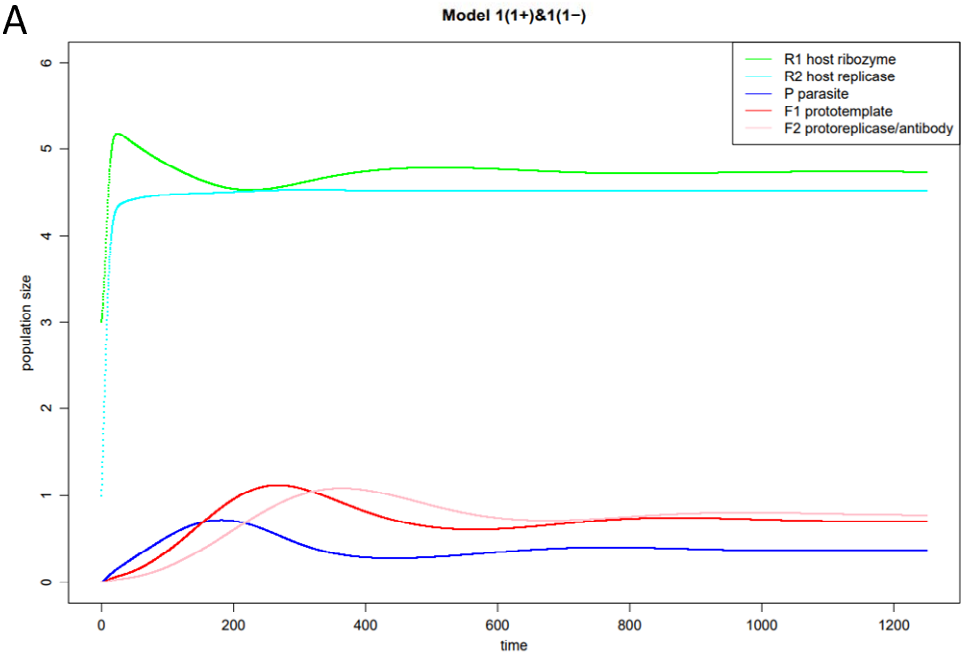

Figure 4. Transition from hyperparasitism to antiparasitism under strong parasite inflow

Fig. 4A Under strong, permanent exposure to complex parasite populations the hyperparasite cycle partially loses its catalytic activity and becomes antiparasitic. The protoreplicase F_2 now continues to function as a partially active catalyst, but also exerts a non-catalytic antiparasitic function (antibody). The parameters in Fig. 4A are set to have these two activities equally represented.

Fig. 4B

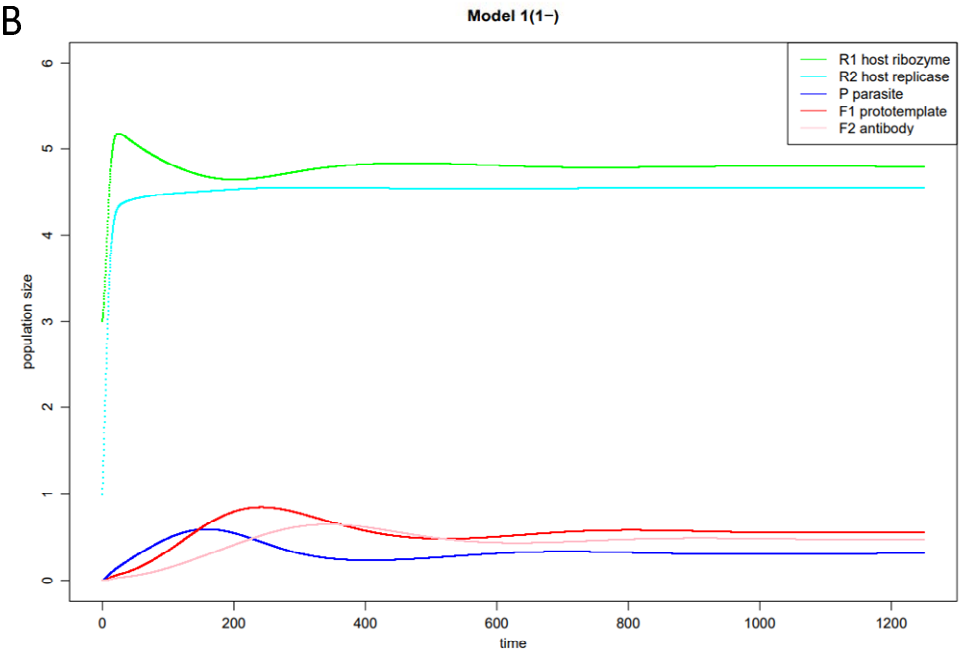

Fig. 4B The parameters are now set to completely shift to antiparasitism (fully degenerated hyperparasite cycle, completely non-catalytic). Antiparasitism has fully replaced hyperparasitism, and the habitat is still stabilized, yet with stronger fluctuations.

Fig. 5

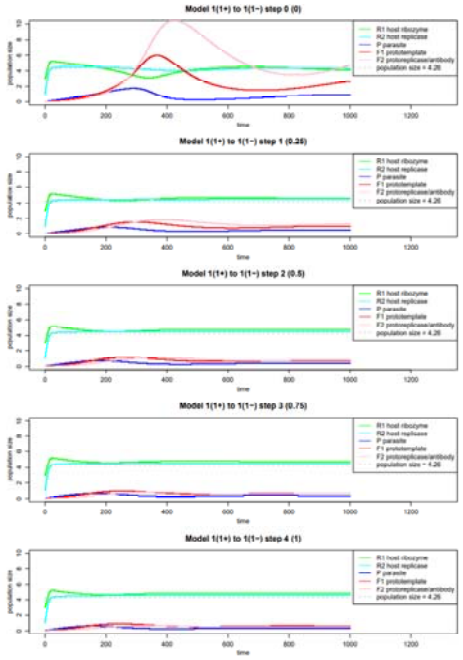

Figure 5. Gradual transition from hyperparasitism to antiparasitism in four steps

The stepwise transition from pure hyperparasitism (top panel) to pure antiparasitism (botton panel) in four steps is shown (0, 0.25. 0.5, 0.75 and 1).