Selection dynamics in transient compartmentalization

Alex Blokhuis,[1] David Lacoste,[1] Philippe Nghe,[2] and Luca Peliti[3,*]

[1] *Gulliver Laboratory, UMR CNRS 7083, PSL Research University, ESPCI, 10 rue Vauquelin, F-75231 Paris Cedex 05, France*
[2] *Laboratory of Biochemistry, PSL Research University, ESPCI, 10 rue Vauquelin, F-75231 Paris Cedex 05, France*
[3] *SMRI, 00058 Santa Marinella (RM), Italy*

(Dated: March 14, 2018)

Transient compartments have been recently shown to be able to maintain functional replicators in the context of prebiotic studies. Here, we show that a broad class of selection dynamics is able to achieve this goal. We identify two key parameters, the relative amplification of non-active replicators (parasites) and the size of compartments. These parameters account for competition and diversity, and the results are relevant to similar multilevel selection problems, such as those found in virus-host ecology and trait group selection.

PACS numbers: 05.40.-a, 87.14.G-, 87.23.Kg

A central issue in origin of life studies is to explain how replicating functional molecules could have appeared and evolved towards higher complexity [1]. In 1965, Spiegelman showed experimentally that RNA could be replicated by an enzyme called $Q\beta$ RNA replicase, in the presence of free nucleotides and salt. Interestingly, he noticed that as the process is repeated, shorter and shorter RNA polymers appear, which he called parasites. Typically, these parasites are non-functional molecules which replicate faster than the RNA polymers introduced at the beginning of the experiment and which for this reason tend to dominate. Eventually, a polymer of only 218 bases remained out of the original chain of 4500 bases, which became known as Spiegelman's monster. In 1971, Eigen conceptualized this observation by showing that for a given accuracy of replication and relative fitness of parasites, there is a maximal genome length that can be maintained without errors [2]. This result led to the following paradox: to be a functional replicator, a molecule must be long enough. However, if it is long, it can not be maintained since it will quickly be overtaken by parasites. Many works attempted to address the puzzle as reviewed in Ref. [3]. In some recent studies, spatial clustering was found to promote the survival of cooperating replicators [4–6]. This kind of observation is compatible with early theoretical views [7, 8], that compartmentalization could allow parasites to be controlled.

Small compartments are ideal for prebiotic scenarios, because they function as micro-reactors where chemical reactions are facilitated. Oparin imagined liquid-like compartments called coacervates, which could play a central role in the origin of life [9]. Although experimental verification of the prebiotic relevance of coacervates or other sorts of protocells remained scarce for many years, the idea has resurfaced recently in various systems of biological interest [10, 11]. An important aspect of the original Oparin scenario which has not been addressed in these studies is the possibility of a transient nature of the compartmentalization. In the present paper, we introduce a general class of multilevel selection with transient compartimentalization. This class includes several scenarios [12–16] for the origin of life and a recent experiment, in which small droplets containing RNA in a microfluidic device [17] were used as compartments. In this experiment, cycles of transient compartmentalization prevent the takeover by parasitic mutants. Cycles consist of the following steps: (i) inoculation, in which droplets are inoculated with a mixture of RNA molecules containing active ribozymes and inactive parasites, (ii) maturation, in which RNA is replicated by $Q\beta$ replicase, (iii) selection, in which compartments with a preferred value of the catalytic activity are selected, (iv) pooling, in which the content of the selected compartments is pooled. This protocol does not correspond to that of the Stochastic Corrector model [7] because of step (iv), which removes the separation between individual compartments. Instead, it corresponds to group selection with a pooling phase [18].

The absence of parasite takeover was successfully explained in ref. [17] by a theoretical model which described the appearance of parasites within a given lineage as a result of mutations during the replication process. In this work we wish to account for these observations in a more general sense. We show that the value of the mutation rate does not play an essential role as long as it is small [19], and that the entire shape of the selection function is not needed to describe the fate of

*luca@peliti.org

the system.

Let us consider an infinite population of compartments. Each compartment is initially seeded with n replicating molecules, where n is a random variable, Poisson distributed with average equal to λ. In addition, each compartment also contains a large and constant numbers of enzymes, $n_{Q\beta}$ and of activated nucleotides n_u. Among the n replicating molecules, m are ribozymes, and the remaining $n-m$ are parasites. Let x be the initial fraction of ribozymes and $1-x$ that of parasites. After this inoculation phase, compartments evolve by letting the total number of molecules grow by consuming activated nucleotides.

In practice, the time of incubation of the compartments is fixed and longer than the time after which activated nucleotides become exhausted. The kinetics is initially exponential because the synthesis of RNA is autocatalytic at low concentration of templates. Therefore, the average number $\bar{m}$ of ribozymes and $\bar{y}$ of parasites grow according to

$$\begin{aligned} \bar{m} &= m\exp(\alpha T), \\ \bar{y} &= (n-m)\exp(\gamma T), \end{aligned} \tag{1}$$

where T denotes the time and α (resp. γ) denote the average growth rate of the ribozymes (resp. parasites) during this exponential growth phase. The relevant quantity for this dynamics is the ratio of the number of daughters of one parasite molecule and that of the daughters of one ribozyme molecule: $\Lambda = \exp((\gamma-\alpha)T))$. Note that $\Lambda > 1$ since $\gamma > \alpha$. This exponential growth phase (maturation phase) ends, when the total number of templates $N = \bar{m}+\bar{y}$ reaches the constant value $n_{Q\beta}$ which is the same for all compartments. After this point, the kinetics switches to a linear one, because enzymes rather than templates are limiting [20]. Importantly, during this linear regime the ratio of ribozymes and parasites

$$\bar{x}(n,m) = \frac{\bar{m}}{N} = \frac{m}{n\Lambda-(\Lambda-1)m} \tag{2}$$

does not change. Apart from neglecting very small fluctuations in $n_{Q\beta}$ and n_u, our assumption that N is constant means that the effects of fluctuations of growth rates of both species and the effect of a possible dependence of Λ on m and n are not considered. These two stochastic effects have been modeled in detail in the Supplemental Materials [19]. In the end, we find that they do not alter significantly the predictions of the present deterministic model for the conditions of the experiment.

Two types of parasites can appear: *hard parasites* are formed when the replicase overlooks or skips a large part of the sequence of the ribozyme during replication. The resulting polymers are significantly shorter than that of the ribozyme and will therefore replicate much faster. Based on the experiments of [17], we estimated Λ to be in a range from 10 to about 470, as explained in the Supplemental Materials [19]. In contrast, if the replicase makes errors but keeps overall the length of the polymers unchanged, then the replication time is essentially unaffected. In that case, one speaks of *soft parasites*, and the corresponding Λ is close to unity. It is important to appreciate that the distinction between *hard* or *soft parasites* is not only a matter of replication rates, because Λ also contains the time T, so depending on both parameters, parasites could be classified as either *hard* or *soft*.

The compartments are then selected according to a selection function $f(\bar{x}) \geq 0$. A specific form which is compatible with [17] is the sigmoid function

$$f(\bar{x}) = 0.5\left(1+\tanh\left(\frac{\bar{x}-x_{th}}{x_w}\right)\right), \tag{3}$$

with $x_{th} = 0.25$ and $x_w = 0.1$. Note that this function takes a small but non-zero value for $\bar{x} = 0$, namely $0.5(1-\tanh(x_{th}/x_w)) = 0.0067$, which represents the fitness of a pure parasite compartment. This is in contrast with the linear selection function chosen in a recent study of a similar system [21].

After the selection phase, the resulting products are pooled and the process is restarted with newly formed compartments. We wish to evaluate the steady-state ratio x of ribozymes when many rounds of the process have taken place. The probability distribution of the initial condition (n,m) is given by

$$P_\lambda(n,x,m) = \text{Poisson}(\lambda,n)B_m(n,x), \tag{4}$$

where $B_m(n,x)$ is the Binomial distribution for $m \in \{0,\ldots,n\}$ of parameter $x \in [0,1]$. The average of $\bar{x}$ after the selection step is given by

$$x'(\lambda,x) = \frac{\sum_{n,m}\bar{x}(n,m)f(\bar{x}(n,m))P_\lambda(n,x,m)}{\sum_{n,m}f(\bar{x}(n,m))P_\lambda(n,x,m)}. \tag{5}$$

The steady-state value of x is the stable solution of

$$x = x'(\lambda,x). \tag{6}$$

It is easier to evaluate $\Delta x = x'(\lambda,x)-x$ as a function of λ. The steady-state value corresponds to the line $\Delta x = 0$ separating negative values above from positive values below as shown in Fig. 1.

We construct a phase diagram in the (λ,Λ) plane, by numerically evaluating the bounds of stability of the fixed point $x = 0$ from the condition:

$$\left.\frac{\partial x'}{\partial x}\right|_{x=0} = 1, \tag{7}$$

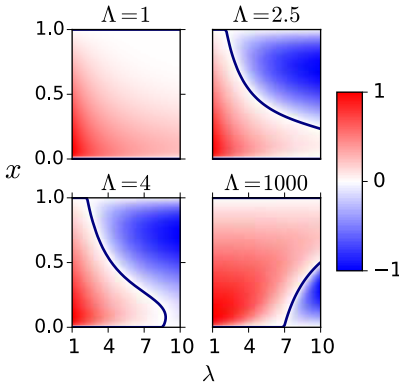

FIG. 1. Contour plots of Δx for four values of $\Lambda = 1, 2.5, 4$ and 1000 in the plane (x, λ), with red (resp. blue) regions corresponding to $\Delta x > 0$ (resp. $\Delta x < 0$).

and similarly for the other fixed point $x = 1$. The resulting phase diagram, as shown in Fig. 2, shows four distinct phases. In the orange (resp. light blue P region) region R, the only stable fixed point is $x = 1$ (resp. $x = 0$). In the green region, $x = 0$ and $x = 1$ are both stable fixed points. The system converges towards one fixed point or the other depending on the initial condition: for this reason we call this region B for bistable. In the violet region, $x = 0$ and $x = 1$ are both unstable fixed points, but there exists a third stable fixed point x^* with $0 < x^* < 1$. We call this a coexistence region (C). All of these phases can be seen in Fig. 1. In the Supplemental Materials [19], we discuss other aspects of the phase behavior which are not captured by this treatment. We also show there that many features of this phase diagram remain if a linear selection function is used instead of Eq. (3)

It is interesting to analyze separately some specific limits for which the asymptotes of the phase diagram can be computed exactly. Let us consider

- $\lambda \gg 1$: bulk behavior
- $\Lambda \gg 1$: *hard* parasites
- Λ close to 1: *soft* parasites

For large λ, we can neglect the fluctuations of n, i.e. the total number of replicating molecules (ribozymes plus parasites) in the seeded compartment. Indeed, n is Poisson distributed with parameter λ, therefore $\mathrm{Var}(n)/\lambda^2 = 1/\lambda \ll 1$. For large λ, Λ close to 1 and x close to 1 (resp. 0), the most abundant compartments verify $m = n$ or $m = n-1$ (resp. $m = 0$ or $m = 1$). By considering only these compartments in the recursion relation [19], one finds that the condition of stability of the fixed point $x = 0$ leads to

$$\Lambda = 1 + \frac{f'(0)}{f(0)\lambda} + O\Big(\frac{1}{\lambda^2}\Big), \tag{8}$$

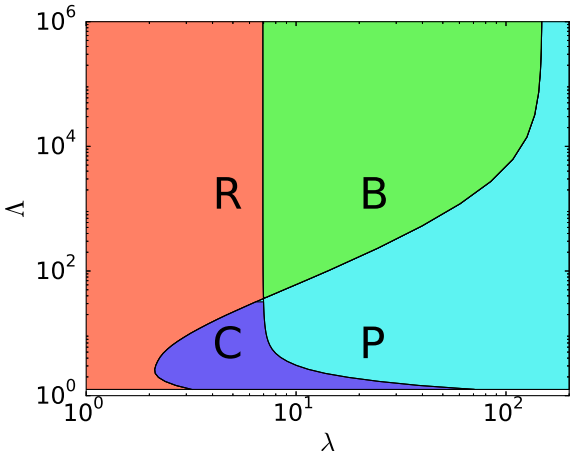

FIG. 2. Phase diagram of the transient compartmentalization dynamics with the selection function of Eq. (3) in the (λ, Λ) plane. The phases are: R: pure Ribozyme, B: Bistable, C: Coexistence, P: pure Parasite.

for an arbitrary selection function and $\Lambda \simeq 1 + 19.86/\lambda$ for the selection function of Eq. (3). This equation indeed characterizes the separation between the parasite and coexistence regime at large λ in Fig. 2. A similar equation is found for the fixed point at $x = 1$

$$\Lambda = 1 + \frac{f'(1)}{f(1)\lambda} + O\Big(\frac{1}{\lambda^2}\Big), \tag{9}$$

yielding $\Lambda \simeq 1 + 6.12\,10^{-6}/\lambda$ for this selection function for the separation between ribozyme and coexistence regions. For Λ close enough to 1, we have a ribozyme phase. The asymptotes given by (8) and (9) border the coexistence region in Fig. 2. This supports the observation that *soft* parasites can coexist with ribozymes.

Let us now study the *hard* parasite limit, namely $\Lambda \gg 1$, and finite λ. In this regime, we only need to consider three types of compartments: compartments made of pure ribozymes, such that $m = n \neq 0$, compartments containing parasites, and empty compartments, i.e. such that $n = 0$. One can introduce three inoculation probabilities for these cases p_{ribo}, p_{para}, and p_{zero}. Using Eq. (4), one finds

$$\begin{aligned} p_{ribo} &= \sum_{n=1}^{\infty} \frac{x^n \lambda^n}{n!} e^{-\lambda} = (e^{\lambda x} - 1)e^{-\lambda}, \\ p_{zero} &= e^{-\lambda}, \\ p_{para} &= 1 - p_{ribo} - p_{zero} = 1 - e^{\lambda(x-1)}. \end{aligned} \tag{10}$$

Assuming that in compartments containing parasites they will overwhelm the ribozymes, and in-

serting these values in (5), we find

$$x' = \frac{p_{ribo} f(1)}{p_{ribo} f(1) + p_{para} f(0)}. \quad (11)$$

Evaluating the fixed-point stability of $x = 1$ using (7), we find that the boundary value of λ satisfies

$$\lambda f(0) e^{\lambda} = (e^{\lambda} - 1) f(1), \quad (12)$$

for an arbitrary selection function. A similar calculation at the fixed point $x = 0$ leads to the other vertical separation line given by

$$\lambda f(1) = (e^{\lambda} - 1) f(0). \quad (13)$$

The solution of Eq. (12) (resp. Eq. (13)) is $\lambda \simeq 149.41$ (resp. $\lambda \simeq 6.95$) which compare well with the vertical separation lines in Fig. 2.

In ref. [17] a comparison was made of the system behavior as a function of the number of selection rounds in three possible protocols: (i) No compartments (bulk behavior), (ii) compartments with no selection, (iii) compartments with selection. Such a comparison based on our theoretical model is shown in Fig. 3 for parameter values corresponding to the coexistence region of Fig. 2. As

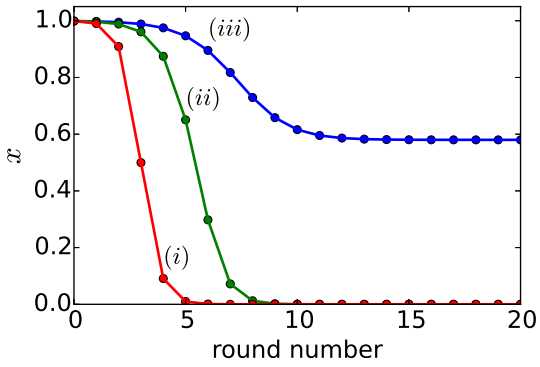

FIG. 3. Evolution of the average ribozyme fraction x as function of the number of rounds for the three protocols, namely (i) No compartments (bulk behavior), (ii) compartments with no selection, (iii) compartments with selection. We choose $\lambda = 5$ and $\Lambda = 10$, corresponding to the coexistence region of Fig. 2.

expected, the fraction of ribozymes decreases towards zero rapidly in case (i), and somewhat less quickly in case (ii). Only in case (iii) is it possible to maintain a non-zero ribozyme fraction on long times. It is indeed observed that the ribozyme fraction eventually vanishes for protocols (i) and (ii) in the experiment of Ref. [17]. In case (iii), a decrease of the ribozyme fraction is observed. The last two points of figure 2C (top panel) in this reference are an indication that the system may eventually reach ribozyme-parasite coexistence in this regime.

In figure 4 we show the behavior of the distribution of the ribozyme fraction after the growth phase, i.e. $\bar{x}(n, m)$ (defined in Eq. (2)) as a function of round number. The parameters are $\Lambda = 5$ and $\lambda = 10$, corresponding to the parasite region, where the final state of the system is $x = 0$, and the initial condition is $x = 0.999$. Note that the distribution of $\bar{x}(n, m)$ is discrete, since many values are not accessible in the allowed range of n and m. At $t = 0$, it exhibits a sharp peak near $\bar{x} = 1$ coexisting with a broad peak at small values of $\bar{x}$. As time proceeds, the weight of the distribution shifts to the peak at small values of $\bar{x}$, since in this case selection is not sufficiently strong to favor the peak near $\bar{x} = 1$ and parasites eventually take over.

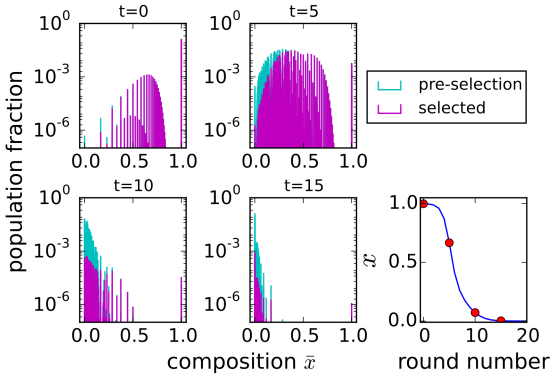

FIG. 4. Evolution of the distributions of ribozyme fraction $\bar{x}(n, m)$ before and after selection at different times. The chosen times are shown as red circles in the lower right panel, which represents the evolution of the average fraction x as a function of the number of selection rounds.

In conclusion, we captured the behavior of transient compartmentalization with a model containing only two parameters, which remarkably suffices to capture the main features of the transient compartimentalization experiment [17]. The model predictions are summarized in a phase diagram, which has been derived for an arbitrary selection function.

Given its basic ingredients, the competition between a host and its parasite, and the diversity generated by small size compartments, which is required for selection to be efficient [22], the model has broad applicability. It is relevant for phage-bacteria ecology problems, since phages experience a similar life cycle of transient replication in cellular compartments during infection [23]. More generally it is relevant for the issue of cooperation between producers and non-producers [24].

Our work also shows that selection is able to purge the parasites even when compartments are

transient. Examples of transient compartments are protocells with [13] or without membranes [9] and rock pores [25], which have all been considered in scenarios for the origins of life. The fact that selection here occurs at the group level while heredity is carried out by replicating molecules at a different molecular level underlines the importance of multilevel selection in order to explain major evolutionary transitions [26].

A.B. was supported by the Agence Nationale de la Recherche (ANR-10-IDEX-0001-02, IRIS OCAV). L.P. acknowledges support from a chair of the Labex CelTisPhysBio (ANR-10-LBX-0038). He would like to thank ESPCI and its director, J.-F. Joanny, for a most pleasant hospitality.

[1] P. G. Higgs and N. Lehman, Nat. Rev. Genet. **16**, 7 (2015).
[2] M. Eigen, Naturwissenschaften **58**, 465 (1971).
[3] N. Takeuchi and P. Hogeweg, Physics of Life Reviews **9**, 219 (2012).
[4] S. R. Levin and S. A. West, Proc. R. Soc. B **284** (2017).
[5] A. S. Tupper and P. G. Higgs, J. Theor. Biol. **428**, 34 (2017).
[6] Y. E. Kim and P. G. Higgs, PLoS Comput Biol **11**, 34 (2016).
[7] E. Szathmáry and L. Demeter, J. Theor. Biol **128**, 463 (1987).
[8] J. Maynard Smith and E. Szathmáry, *The Major Transitions in Evolution* (Freeman, Oxford, 1995).
[9] A. I. Oparin, *Origin of Life* (Dover, 1952).
[10] D. Zwicker, R. Seyboldt, C. A. Weber, A. A. Hyman, and F. Jülicher, Nat. Phys. **13**, 408 (2017).
[11] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Science **324**, 1729 (2009).
[12] E. V. Koonin and W. Martin, Trends Genet. **21**, 647 (2005).
[13] P. L. Luisi, P. Walde, and T. Oberholzer, Curr. Op. Coll. Int. Sci. **4**, 33 (1999).
[14] P. Szabó, I. Scheuring, T. Czárán, and E. Szathmáry, Nature **420**, 340 (2002).
[15] B. Damer and D. Deamer, Life **5**, 872 (2015).
[16] P. Baaske, F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell, and D. Braun, Proc. Natl. Acad. Sci. U.S.A. **104**, 9346 (2007).
[17] S. Matsumura, A. Kun, M. Ryckelynck, F. Coldren, A. Szilágyi, F. Jossinet, C. Rick, P. Nghe, E. Szathmáry, and A. D. Griffiths, Science **354**, 1293 (2016).
[18] D. S. Wilson, Proc. Natl. Acad. Sci. USA **72**, 143 (1975).
[19] See Supplemental Material for the determination of the parameter Λ from the data of ref. [14], a discussion of a stochastic version of the present model and of other details concerning the phase diagram.
[20] S. Spiegelman, I. Haruna, I. B. Holland, G. Beaudreau, and D. Mills, Proc. Natl. Acad. Sci. USA **54**, 919 (1965).
[21] A. S. Zadorin and Y. Rondelez, ArXiv e-prints (2017), arXiv:1707.07461 [q-bio.PE] .
[22] R. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).
[23] K. Sneppen, *Models of life* (Cambridge University Press, 2014).
[24] J. S. Chuang, O. Rivoire, and S. Leibler, Science **323**, 272 (2009).
[25] M. Kreysing, L. Keil, S. Lanzmich, and D. Braun, Nature chemistry **7**, 203 (2015).
[26] N. Takeuchi, P. Hogeweg, and K. Kaneko, Nature Communications **8**, 250 (2017).