Huge progeny production during the transient of a quasi-species model of viral infection, reproduction and mutation

José A. Cuesta

Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28911 Leganés, Madrid, Spain

Abstract

Eigen's quasi-species model describes viruses as ensembles of different mutants of a high fitness "master" genotype. Mutants are assumed to have lower fitness than the master type, yet they coexist with it forming the quasi-species. When the mutation rate is sufficiently high, the master type no longer survives and gets replaced by a wide range of mutant types, thus destroying the quasi-species. It is the so-called "error catastrophe". But natural selection acts on phenotypes, not genotypes, and huge amounts of genotypes yield the same phenotype. An important consequence of this is the appearance of beneficial mutations which increase the fitness of mutants. A model has been recently proposed to describe quasi-species in the presence of beneficial mutations. This model lacks the error catastrophe of Eigen's model and predicts a steady state in which the viral population grows exponentially. Extinction can only occur if the infectivity of the quasi-species is so low that this exponential is negative. In this work I investigate the transient of this model when infection is started from a small amount of low fitness virions. I prove that, beyond an initial regime where viral population decreases (and can go extinct), the growth of the population is super-exponential. Hence this population quickly becomes so huge that selection due to lack of host cells to be infected begins to act before the steady state is reached. This result suggests that viral infection may widespread before the virus has developed its optimal form.

Keywords: evolution, quasi-species, replicator-mutator, population dynamics
2000 MSC: 92D15, 92D25, 92D30, 05A15

1. Introduction

It seems that an unavoidable consequence of the increase in complexity of a system is the appearance of parasites. These are entities able to exploit backdoors, bypasses, holes... of the system for their own benefit, sometimes even at a cost for the system. We see a huge variety of these parasites in biology, ranging from viruses to humans. Society, in fact, is one of those complex systems amenable to exploitation by free-riders (the paradigm of the Public Goods game [1] is but one prominent acknowledgment of the existence of this social parasitism). More recently, the widespread use of computers and the arrival of Internet has made us witness the

Email address: cuesta@math.uc3m.es (José A. Cuesta)
URL: http://gisc.uc3m.es/~cuesta (José A. Cuesta)

emergence and proliferation of computer viruses, trojans, worms, spam, phising, and all kinds of forms of parasitism, which f ood the web using the same mechanisms aimed at allowing the transmission of information. Apparently, whenever a complex mechanism emerges, it is soon invaded by its specif c parasites.

Parasites need not be complex: on the contrary, by being very specif c to a particular mechanism, they are able to do their job with very simple mechanisms. Paradigmatic among parasites for their extreme simplicity are viruses. Their success is such that they are the most abundant life forms on Earth [2]. Their existence is an unavoidable outcome of the very evolutionary process. In fact, the most common strategy of RNA viruses is to have a very high reproductive rate which yields a wide variety of mutants [3]. This ensures their fast adaptation to almost any change.

One of the most important challenges in current medical research is how to f ght viruses, and one of the most studied strategies is the design of therapies able to induce viral extinction. Increasing the mutation rate has been successful, at least in experiments *in vitro,* but there is no consensus as to why the virus loses infectivity at high mutation rates [4, 5, 6, 7]. The pioneering work of Eigen [8] explains viral extinction through a mechanism known as error threshold. According to it, the progeny loses its identity if the mutation rate grows above a given value, which is inversely proportional to the length of the replicating molecule —hence putting an upper bound to the complexity of viruses. This classical theory is currently questioned. The current state of the art of the evolutionary paradigm contradicts some of the basic assumptions of Eigen's theory, crucial for the existence of the error threshold. Alternative mechanisms may lead to viral extinction for reasons other than this hypothetical error threshold (like the presence of defective forms of the virus [9, 10], the competition induced by geometrical constraints [11, 12, 13, 14, 15], etc.).

Models of viral evolution need to make simplifying assumptions, and real virus behavior often deviates substantially from their predictions [16]. Current quasi-species models assume high mutation rates that give rise to heterogeneous populations. This is consistent with experimental observations. However, one common approximation is to consider that all new mutations have a deleterious effect on f tness, thus neglecting benef cial and neutral mutations. This is true if, as the theory assumes, there is a unique master sequence of high f tness. But we now know that the genotype-phenotype map is extremely redundant, and that a huge amount of different sequences —forming so-called neutral networks [17]— yield phenotypes that perform equally well. The increase in the rate of benef cial and neutral mutations that this effect brings about invalidates the classical theory of the error threshold [7] and calls for alternative models of viral evolution and extinction.

The aim of this paper is to explore one such model, introduced by Manrubia et al. [18], with special focus on its transient behavior.

2. Quasi-species equation

Evolution is a result of the simultaneous action of three processes: replication, mutation and selection. Any set of agents undergoing these three processes evolve —in the direction determined by selection— regardless of whether they are biological entities, computer programs, cultural traits, etc.

Replication is the ability of some agents to produce identical copies of themselves. Replication is normally a stochastic process, characterized by a probability distribution $p(k)$, $k = 0, 1, 2, \ldots$, representing the probability that after a replication event —however we def ne it— there are k replicas of the parent agent (including itself).

The replication process is usually imperfect. Most often errors in making copies yield invalid individuals (unable to produce further copies); however sometimes these errors produce valid individuals albeit of a different type (or species). These kind of altered replications are referred to as mutations. Mutations create new species and maintain variability within populations. New species may have modified replicative abilities, and so a probability distribution $p_i(\mathbf{k})$ must be introduced for each species i, where $\mathbf{k} = (k_1, k_2, \ldots, k_s)$ is a vector denoting the number of offspring of any of the s possible resulting species that an individual of species i gives rise to.

The replication with mutation of an individual of any of the valid species generates a Markov process in discrete time known as multi-type branching process [19]. The variable characterizing this process is the population of each species at generation t, $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \ldots, Z_s(t))$. The mean value of this variable $\mathbf{n}(t) = E[\mathbf{Z}(t)]$ has the simple evolution equation

$$\mathbf{n}(t + 1) = \mathbf{n}(t)W, \tag{1}$$

where $W = (w_{ij})$ is the replication-mutation matrix. The number $r_i = \sum_j w_{ij}$ denotes the average number of offspring that an individual of species i produces in a replication event, and $q_{ij} = w_{ij}/r_i$ is the probability that one of this offspring mutates to species j. So introducing stochastic matrix $Q = (q_{ij})$ (mutation matrix) and the diagonal matrix $R = (r_i \delta_{ij})$ (replication matrix), we can factorize $W = RQ$, thus separating the effect of replication and mutation in the evolution of $\mathbf{n}(t)$.

The asymptotic behavior of this equation is given by $\mathbf{n}(t) = \lambda^t \mathbf{u}$, where λ is the largest eigenvalue of W and $\mathbf{u}$ a (positive) eigenvector of its corresponding eigenspace. Population grows exponentially if $\lambda > 1$, or vanishes exponentially if $\lambda < 1$.[1]

We have not considered selection yet. Selection is induced by the environment, usually through a finite availability of resources for replication. Selection thus acts on the specific replicative ability of each species —modifying the values of r_i. When scarcity of resources affects species equally, all values of r_i are affected equally. In that case, what determines the fate of each species is its asymptotic fraction within the population. At generation t the fractions of population of each species is given by the vector $\mathbf{x}(t) = \mathbf{n}(t)/\mathbf{n}(t) \cdot \mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)$. Equation (1) then becomes

$$\mathbf{x}(t + 1) = \phi(t)^{-1}\mathbf{x}(t)W, \qquad \phi(t) = \mathbf{x}(t)W\mathbf{1}^\mathsf{T} = \sum_i r_i x_i(t), \tag{2}$$

where we have used the factorization $W = RQ$ and the fact that Q is stochastic (hence $Q\mathbf{1}^\mathsf{T} = \mathbf{1}^\mathsf{T}$). Function $\phi(t)$ represents the mean replicative ability of the population at generation t. Equation (2) is referred to as the *quasispecies equation*.

The steady state of equation (2) is obtained by solving the eigenvalue problem $\mathbf{x}W = \phi\mathbf{x}$, under the constraint $\sum_i x_i = 1$, $x_i > 0$, $i = 1, \ldots, s$. If Q is an irreducible matrix ϕ and $\mathbf{x}$ are respectively the largest eigenvalue and its corresponding (unique) normalized left eigenvector of matrix W [20].

3. Error catastrophe

Eigen proposed the quasi-species equation as a model for the evolution of prebiotic replicators [8] which, in the absence of correction mechanisms, had a high mutation rate and accordingly

---

[1]If $\lambda = 1$ the process is "critical", and it can be proven to go extinct in finite time with probability one [19].

a short length. However it has become a paradigm of viral evolution even for much longer sequences (RNA, DNA, proteins. . .) [5, 21]. To envisage Eigen's idea we can think of a space of L long sequences, labelled $i = 0, 1, \ldots, s$. Each position of these sequences can be occupied by any element of a given set of them (DNA or RNA bases, alleles of genes, aminoacids. . .). Let us assume that this set contains a elements ($a = 4$ for bases, $a = 20$ for aminoacids. . .). Mutations are point-like, i.e., substitutions of the element at a single position by any other in the set. Thus sequences ACGGCA and AGGGCA are reached from each other by a mutation, whereas ACGGCA and AGGGCC are two mutations apart. Any offspring of the replicated sequence will carry a point mutation with probability $0 < \mu \ll 1$. The sequence labeled as 0 (master sequence) is assumed to have a higher replicative ability (henceforth *f tness)* than any other sequence. For simplicity, all sequences are assigned f tness 1 whereas the master sequence has f tness $f > 1$. We shall denote the fraction of population of the master sequence by x.

An important assumption in Eigen's model is that backward mutations that recover the master sequence are neglected. This is a reasonable assumption considering that sequences in nature tend to be very long. The master sequence is recovered with probability $(\mu/D)^h$, where h is the Hamming distance (number of different positions) between the given sequence and the master sequence, and $D = (a - 1)L$ is the number of point mutants of an L long sequence.

Under the above assumptions the quasi-species equation (2) reads

$$x[f(1 - \mu D) + \epsilon] = x\phi, \qquad \phi = 1 + (r - 1)x, \tag{3}$$

where ϵ contains those backwards mutation that the theory neglects. This equation predicts

$$x \approx \begin{cases} 1 - \frac{r}{r-1}\mu D & \text{if } \mu D < 1 - \frac{1}{r}, \\ 0 & \text{if } \mu D > 1 - \frac{1}{r}, \end{cases} \tag{4}$$

in other words, if the mutation rate is above a threshold (which decreases as L^{-1}), the master sequence accumulates so many mutations that it gets lost in a cloud of mutants. This transition is known as the *error catastrophe* and has provided a line of research to f nd a therapy against viral infections based on increasing μ through the addition of mutagens [22].

4. Phenotype vs. genotype

But Eigen's model is fundamentally f aw in the assuming the existence of a single master sequence or genotype. Biology is extremely redundant. DNA codes for proteins using a (nearly) universal genetic code based on triplets of bases or codons. Each codon codes for an aminoacid. But the 64 possible codons only code for 20 aminoacids plus a STOP signal. In redundant aminoacids, typically the third base is irrelevant or nearly so. This means that many mutations changing a base pair in the DNA sequence remain silent when transcribed into proteins. On their side, proteins fold in a three-dimensional structure which determines their function. And only a few aminoacids at selected positions are key to this folding. So the replacement of many of them leaves the protein structure (hence its function) intact. Evolution can only act on the macroscopic features of living beings (their phenotype), which are blind to a huge amount of mutations. In other words, the mapping from genotypes into phenotypes is from very many to one. The existence of a master sequence is therefore an entelechy. At most we can only speak of a master phenotype.

The distribution of genotypes corresponding to a given phenotype on genotype space is a rather complicated one. Basically they form so-called neutral networks [17], i.e., connected components of the mutation graph through which sequences can be changed by successive mutations without ever changing the phenotype —hence their f tness. The most relevant consequence of the existence of neutral networks is that backwards (or benef cial) mutations are not negligible, because recovering the master phenotype (not genotype) is no more an improbable event. Changing Eigen's model to account for benef cial mutations eliminates the error catastrophe, as we will see in what follows.

5. A model with benef cial mutations

A simple model accounting for the existence of neutral networks has been recently proposed [18]. In this model, viral phenotypes are characterized by their replicative abilities, $r \in \{0, 1, \ldots, R\}$. The only mutations that the model takes into account are those connecting neighboring classes (i.e., the effect of a mutation is a slight increase or decrease in the replicative ability). An offspring undergoes a deleterious mutation from from class r to class $r - 1$ with probability p, and a benef cial mutation from class r to class $r + 1$ with probability q. In general it is assumed that $0 < q \ll p \ll 1$. If we denote $n_r(t)$ the mean number of viral particles in class r at generation t, then

$$n_r(t + 1) = (1 - p - q)r n_r(t) + p(r + 1)n_{r+1}(t) + q(r - 1)n_{r-1}(t), \qquad r = 1, 2, \ldots R - 1,$$
$$n_R(t + 1) = (1 - p)R n_R(t) + q(R - 1)n_{R-1}(t). \tag{5}$$

Here R stands for the maximum replicative ability of the virus. There exists also class $r = 0$, with no replicative ability, whose population is maintained because of deleterious mutations from class $r = 1$. Hence $n_0(t) = p n_1(t)$.

Equations (5) have the form of (1) for $W = RQ$ with

$$Q = \begin{pmatrix} 1 - p - q & q & & & \\ p & 1 - p - q & q & & \\ & \ddots & \ddots & \ddots & \\ & & p & 1 - p - q & q \\ & & & p & 1 - p \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & R \end{pmatrix}. \tag{6}$$

Notice however that Q is only sub-stochastic if we do not include class $r = 0$. This fact may cause the total extinction of the virus. Still, the eigenvalue equation $\phi \mathbf{u} = \mathbf{u}W$ determines the asymptotic behavior of the system $\mathbf{n}(t) \sim \phi^t \mathbf{u}$. Both ϕ and $\mathbf{u}$ are unique because W is irreducible. Vector $\mathbf{u}$ normalized as $\mathbf{u} \cdot \mathbf{1} = 1$ describes the asymptotic fractions of viral particles in each class —even in the case that the virus eventually goes extinct.

For $q = 0$ it is easy to check that $\lambda_r = r(1 - p)$ and $\mathbf{v}_r = (v_{r1}, \ldots, v_{rR})$, with $v_{rk} = \binom{r}{k}(1 - p)^k p^{r-k}$, $r, k = 1, \ldots, R$, are the eigenvalues and left eigenvectors of matrix W, respectively. Since for every p the largest eigenvalue is $\phi = R(1 - p)$, we f nd that $p_c = 1 - R^{-1}$ def nes a transition value such that the virus proliferates for $p < p_c$ but gets extinct for $p > p_c$. This transition is similar to Eigen's error catastrophe, except for the fact that the virus becomes extinct in this case because the lowest f tness class is $r = 0$, unable to infect further cells.

6. Transient and the infinite classes model

As for the case $q = 0$, for $q > 0$ we expect that the largest eigenvalue $\phi = O(R)$, so a model with an infinite number of classes will never reach the asymptotic state. However such a model can be useful to study the initial stages of the transient behavior provided $R \gg 1$ and initially the population has a low fitness $r_0 \ll R$. The reason is that classes above r_0 get populated one by one, so at least for $0 \le t \le R - r_0$ there is no difference between the model with $R < \infty$ and with $R = \infty$. This is the regime I plan to analyze here.

So consider that the first of equations (5) holds for all $r \in \mathbb{N}$ and assume that $n_r(0) = 0$ for all $r > r_0$. Then the generating function

$$G(z, t) \equiv \sum_{r=1}^{\infty} z^r n_r(t) \tag{7}$$

will be a polynomial of degree at most $r_0 + t$. Multiplying (5) by z^r and adding up for all $r \ge 1$ we obtain

$$G(z, t + 1) = \left[p + (1 - p - q)z + qz^2 \right] G_z(z, t) - pn_1(t). \tag{8}$$

(Subindexes in functions are meant to denote partial derivatives.)

Let us now introduce the generating functions

$$N_r(s) \equiv \sum_{t=0}^{\infty} \frac{s^t}{t!} n_r(t), \qquad F(z, s) \equiv \sum_{t=0}^{\infty} \frac{s^t}{t!} G(z, t) = \sum_{r=1}^{\infty} z^r N_r(s). \tag{9}$$

In terms of them equation (8) becomes

$$F_s(z, s) = p\left(1 + \frac{z}{w_-}\right)\left(1 + \frac{z}{w_+}\right) F_z(z, s) - pN_1(s). \tag{10}$$

where

$$w_{\pm} \equiv \frac{1 - p - q \pm \Omega}{2q}, \qquad \Omega \equiv \sqrt{1 - 2(p + q) + (p - q)^2}. \tag{11}$$

The condition for Ω to be real and positive is $\sqrt{p} + \sqrt{q} < 1$. This condition holds whenever $0 < q < p < 1/4$. As $p = 1/4$ is an extremely high mutation rate, we shall take for granted that $\Omega \in \mathbb{R}^+$.

The first order partial differential equation (10) needs to be supplemented with an initial condition for $F(z, s)$. Indeed, if $\{n_r(0)\}_{r \ge 1}$ is the initial condition of the viral populations, then

$$F(z, 0) = G(z, 0) \equiv g(z) = \sum_{r=1}^{\infty} z^r n_r(0). \tag{12}$$

The characteristic curves of equation (10) are given by

$$\left(1 + \frac{z}{w_-}\right)\left(1 + \frac{z}{w_+}\right)^{-1} e^{\Omega s} = \zeta, \tag{13}$$

with ζ an arbitrary constant. We can eliminate z from this equation to get

$$z = z(\zeta, s) = \frac{p(\zeta - E)}{q(w_+ E - w_- \zeta)}, \qquad E \equiv e^{\Omega s}. \tag{14}$$

6

In terms of the variables (ζ, s) and denoting $f(\zeta, s) = F(z(\zeta, s), s)$, equation (10) becomes

$$-p^{-1}f_s(\zeta, s) = N_1(s), \qquad f(\zeta, 0) = g(z(\zeta, 0)) = g\left(\frac{p(\zeta - 1)}{q(w_+ - w_-\zeta)}\right), \tag{15}$$

whose solution is

$$f(\zeta, s) = g\left(\frac{p(\zeta - 1)}{q(w_+ - w_-\zeta)}\right) - p\int_0^s N_1(u)\, du. \tag{16}$$

Substituting (13) into (16) yields

$$\frac{p(\zeta - 1)}{q(w_+ - w_-\zeta)} = \frac{z + w_+\chi(z, s)}{1 - \chi(z, s)} \equiv \psi(z, s), \qquad \chi(z, s) \equiv \frac{w_- + z}{w_+ - w_-}(E - 1). \tag{17}$$

It only remains to determine $N_1(s) = F_z(0, s)$. This can be achieved by imposing $F(0, s) = 0$ in (16), which leads to

$$p\int_0^s N_1(u)\, du = g(\psi(0, s)).$$

Thus the final expression of the generating function $F(z, s)$ is

$$F(z, s) = g(\psi(z, s)) - g(\psi(0, s)). \tag{18}$$

7. Asymptotic behavior of the transient

Setting $z = 1$ in (9) we get $F(1, s) = \sum_{t=0}^{\infty} n(t)s^t/t!$, the generating function of the total population of the virus $n(t) = \sum_{r=1}^{\infty} n_r(t)$. From (18), $F(1, s) = g(\psi(1, s)) - g(\psi(0, s))$, where

$$\psi(1, s) = \frac{1 + w_+\chi(1, s)}{1 - \chi(1, s)}, \qquad \chi(1, s) = \frac{w_- + 1}{w_+ - w_-}\left(e^{\Omega s} - 1\right), \tag{19}$$

$$\psi(0, s) = \frac{w_+\chi(0, s)}{1 - \chi(0, s)}, \qquad \chi(0, s) = \frac{w_-}{w_+ - w_-}\left(e^{\Omega s} - 1\right). \tag{20}$$

Let us assume for simplicity that $g(z) = z^r$, i.e., at time $t = 0$ only a single viral particle of class r is present. We can infer the asymptotic behavior of $n(t)$ from the singularities of $F(1, s)$ [23]. There are two sets of singularities: $s_0 + i2\pi n_0/\Omega$ and $s_1 + i2\pi n_1/\Omega$, with $n_0, n_1 \in \mathbb{Z}$, which are the solutions to $\chi(0, s) = 1$ and $\chi(1, s) = 1$, respectively. In each set, the singularity on the real axis is the one with smallest modulus, so we shall ignore the remaining ones. Denote $E_0 = e^{\Omega s_0}$ and $E_1 = e^{\Omega s_1}$; then

$$E_0 = 1 + \frac{w_+ - w_-}{w_-} = \frac{w_+}{w_-} \qquad E_1 = 1 + \frac{w_+ - w_-}{1 + w_-} = \frac{1 + w_+}{1 + w_-}. \tag{21}$$

But

$$\frac{E_0}{E_1} = \frac{w_+(1 + w_-)}{w_-(1 + w_+)} = \frac{w_+ + p/q}{w_- + p/q} \geq 1,$$

because $w_+ \geq w_-$ for all $p, q \geq 0$ (the inequality is strict if at least one of them is nonzero). Then $s_0 \geq s_1$, so s_1 is the singularity that is closest to the origin. From (21)

$$\Omega s_1 = \log\left(\frac{1 + w_+}{1 + w_-}\right) = \log\left(\frac{(1 + q - p + \Omega)^2}{4q}\right). \tag{22}$$

7

As $\lim_{s \to s_1} \frac{d}{ds}[1 - \chi(1,s)] = -\Omega \frac{w_- + 1}{w_+ - w_-} E_1 = -q(w_+ + 1) \neq 0$, then s_1 is a simple pole of $\psi(1,s)$. As its residue is $-1/q$, then

$$\psi(1,s) \sim \frac{1}{q} \frac{1}{s_1 - s} \quad \text{as} \quad s \to s_1, \tag{23}$$

and therefore

$$F(1,s) \sim \frac{1}{(qs_1)^r} \left(1 - \frac{s}{s_1}\right)^{-r} = \frac{1}{(qs_1)^r} \sum_{t=0}^{\infty} \binom{t+r-1}{t} s_1^{-t} s^t, \qquad s_1 \equiv \Omega^{-1} \log\left(\frac{(1+q-p+\Omega)^2}{4q}\right). \tag{24}$$

From this we obtain the asymptotic behavior when $t \to \infty$ of the total population $n(t)$ as

$$n(t) \sim \frac{1}{q^r} \frac{(t+r-1)!}{(r-1)!} s_1^{-t-r} \sim A_r \left(\frac{t+r-1}{es_1}\right)^{t+r-1/2}, \qquad A_r = \frac{1}{(r-1)!q^r} \sqrt{\frac{2\pi e}{s_1}}. \tag{25}$$

8. Discussion

We have analyzed the transient behavior of Manrubia et al.'s model (5) for a very large number of classes ($R \gg 1$), by transforming it into an infinitely many class model. Although an explicit solution cannot be found, I have obtained the generating function associated to the vector of class populations. The singularities of this function provide the time asymptotic behavior of the total population of the virus, valid as long as the number of generations is smaller than R. Surprisingly we find that viral population grows *super-exponentially,* unlike in the steady state.

Analyzing eq. (25) more closely, we notice that s_1 can have very large values and thus induce an initial decay of the population. However, this decay gets dominated by the factorial $(t+r-1)!$ as soon as $t > t_d \equiv es_1 - r + 1$. During this decay time t_d (which is shorter the larger r) fluctuations of the branching process can lead the virus to extinction. Beyond that interval the virus population starts to recover and grows at a faster than exponential rhythm.

A standard assumption in studies of viral quasi-species evolution is that their population is in the exponential asymptotic state. But if $R \gg 1$ the time to reach this state can be very long (in fact, it requires at least $R - r$ generations to reach the optimal class, let alone to attain a stationary distribution among classes). Before that we have the virus population growing faster than exponential and it is plausible that resources get exhausted during this transient period. This means that selection starts playing a role when the steady distribution is not yet established, leading to a behavior different from what is to be expected in the asymptotic regime. The effects of this phenomenon are as yet unexplored.

Acknowledgements

References

[1] T. Groves, J. Ledyard, Optimal allocation of public goods: A solution to the 'free rider' problem, Econometrica 45 (1977) 783–809.

[2] E. V. Koonin, T. G. Senkevich, V. V. Dolja, The ancient virus world and evolution of cells, Biol. Direct 1 (2006) 29.

[3] S. C. Manrubia, E. Lázaro, Viral evolution, Phys. Life Revs. 3 (2006) 65–92.

[4] C. O. Wilke, Quasispecies theory in the context of population genetics, BMC Evol. Biol. 5 (2005) 44.

[5] J. J. Bull, L. A. Meyers, M. Lachmann, Quasispecies made simple, PLoS Comput. Biol. 1 (2005) 450–460.

[6] N. Takeuchi, P. Hogeweg, Error-threshold exists in landscapes with lethal mutants, BMC Evol. Biol. 7 (2007) 15.

[7] S. C. Manrubia, E. Domingo, E. Lázaro, Pathways to extinction – beyond the error threshold, Phil. Trans. R. Soc. 365 (2010) 1943–1952.

[8] M. Eigen, Self-organization of matter and evolution of biological macromolecules, Naturwissenschaften 58 (1971) 465–523.

[9] A. Grande-Pérez, E. Lázaro, P. Lowenstein, E. Domingo, S. C. Manrubia, Suppression of viral infectivity through lethal defection, Proc. Natl. Acad. Sci. USA 102 (2005) 4448–4452.

[10] J. Iranzo, S. C. Manrubia, Stochastic extinction of viral infectivity through the action of defectors, Europhys. Lett. 85 (2009) 18001.

[11] T. Petermann, P. D. L. Ríos, Cluster approximations for epidemic processes: a systematic description of correlations beyond the pair level, J. Theor. Biol. 229 (2004) 1–11.

[12] A. Barrat, M. Barthélemy, A. Vespignani (Eds.), Dynamical processes on complex networks, Cambridge University Press, Cambridge, 2008.

[13] J. Aguirre, S. C. Manrubia, Effects of spatial competition on the diversity of a quasispecies, Phys. Rev. Lett. 100 (2008) 38106.

[14] C. Cases-González, M. Arribas, E. Domingo, E. Lázaro, Benefcial effects of population bottlenecks in an RNA virus evolving at increased error rate, J. Mol. Biol. 384 (2008) 1120–1129.

[15] J. A. Cuesta, J. Aguirre, J. A. Capitán, S. C. Manrubia, The struggle for space: Viral extinction through competition for cells, 2010. Submitted.

[16] M. Eigen, Error catastrophe and antiviral strategy, Proc. Natl. Acad. Sci. USA 99 (2002) 13374–13376.

[17] S. Gavrilets, Fitness Landscapes and the Origin of Species, Princeton University Press, Princeton, 2004.

[18] S. C. Manrubia, E. Lázaro, J. Pérez-Mercader, C. Escarmís, E. Domingo, Fitness distributions in exponentially growing asexual populations, Phys. Rev. Lett. 90 (2003) 188102.

[19] M. Kimmel, D. E. Axelrod, Branching Processes in Biology, Springer, New York, 2002.

[20] E. Seneta, Non-negative Matrices and Markov Chains, Springer, New York, 2006.

[21] E. Domingo (Ed.), Quasispecies: Concept and Implications for Virology, Springer, Berlin, 2006.

[22] S. Crotty, C. E. Cameron, R. Andino, RNA virus error catastrophe: Direct molecular test by using ribavirin, Proc. Natl. Acad. Sci. USA 98 (2001) 6895–6900.

[23] P. Flajolet, R. Sedgewick, Analytic Combinatorics, Cambridge University Press, Cambridge, 2009.