Stochastic Modeling and Simulation of Viral Evolution

Luiza Guimarães Fabreti1 · Diogo Castro1 · Bruno Gorzoni1 · Luiz Mario Ramos Janini2 · Fernando Antoneli3

Received: 22 December 2017 / Accepted: 3 December 2018 / Published online: 14 December 2018
© Society for Mathematical Biology 2018

Abstract
RNA viruses comprise vast populations of closely related, but highly genetically diverse, entities known as quasispecies. Understanding the mechanisms by which this extreme diversity is generated and maintained is fundamental when approaching viral persistence and pathobiology in infected hosts. In this paper, we access quasispecies theory through a mathematical model based on the theory of multitype branching processes, to better understand the roles of mechanisms resulting in viral diversity, persistence and extinction. We accomplish this understanding by a combination of computational simulations and the theoretical analysis of the model. In order to perform the simulations, we have implemented the mathematical model into a computational platform capable of running simulations and presenting the results in a graphical format in real time. Among other things, we show that the establishment of virus populations may display four distinct regimes from its introduction into new hosts until achieving equilibrium or undergoing extinction. Also, we were able to simulate different fitness distributions representing distinct environments within a host which could either be favorable or hostile to the viral success. We addressed the most used mechanisms for explaining the extinction of RNA virus populations called lethal mutagenesis and mutational meltdown. We were able to demonstrate a correspondence between these two mechanisms implying the existence of a unifying principle leading to the extinction of RNA viruses.

Keywords Viral evolution · Quasispecies theory · Branching process · Lethal mutagenesis · Mutational meltdown · Stochastic simulation

Fernando Antoneli
fernando.antoneli@unifesp.br

1 Programa de Pós-Graduação em Infectologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil
2 Departamentos de Microbiologia, Imunologia, Parasitologia and Medicina, Laboratório de Retrovirologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil
3 Departamento de Informática em Saúde, Laboratório de Biocomplexidade e Genômica Evolutiva, Universidade Federal de São Paulo, São Paulo, SP, Brazil
1 Introduction

Viruses with RNA genomes, the most abundant group of human pathogens (Domingo and Holland 1997), exhibit high mutational rates, fast replicative kinetics, large population sizes, and high genetic diversity. Current evidences also indicate that RNA virus populations consist of a wide and interrelated distribution of variants, which can display complex evolutionary dynamics. The complex evolutionary properties of RNA virus populations feature the modulation of viral phenotypic traits, the interplay between host and viral factors, and other emergent properties (Domingo et al. 1985, 2006). During viral infections, these features allow viral populations to escape from host pressures represented by the actions from the immune system, from vaccines and to develop resistance antiviral drugs. Taken together, these features represent the major obstacle for the success and implementation of effective therapeutic intervention strategies.

In order to describe the evolution of RNA viruses and its relationship with their hosts and antiviral therapies, theoretical models of virus evolution have been developed. These models employ mathematical and computational tools as methodological instruments allowing one to address evolutionary questions from a different perspective than the commonly seen use of modern experimental technologies. This kind of approach allows the implementation of low-cost research projects addressing evolutionary questions that are usually investigated by experimental methods. At a deeper level, they provide a systematic perspective of the biological phenomenon, when viewed as proof-of-concept models (Servedio et al. 2014). Verbal or pictorial models have long been used in evolutionary biology to formulate abstract hypotheses about processes and mechanisms that operate among diverse species and across vast time scales. Used in many fields, proof-of-concept models test the validity of verbal or pictorial models by laying out the underlying assumptions in a mathematical framework.

Eigen (1971), Eigen and Schuster (1979) proposed and analyzed a deterministic model for the evolution of polynucleotides in a dialysis reactor based on a system of ordinary, differential equations called quasispecies model. Subsequently, Demetrius et al. (1985) proposed a stochastic quasispecies model in order to overcome some drawbacks of the deterministic quasispecies model of Eigen and Schuster (1979). The approach of Demetrius et al. (1985) employed very powerful methods based on the theory of stochastic branching processes. This theory, originally developed to deal with the extinction of family names (Watson and Galton 1874), has been applied since the forties to a great variety of physical and biological problems (Harris 1963; Athreya and Ney 1972; Kimmel and Axelrod 2002). On the experimental side, an early study of the RNA phage Qβ reporting that sequence variation in a population was high but approximately stable over time around a consensus sequence gave the initial stimulus to consider the notion of quasispecies in the broader context of RNA viruses (Domingo et al. 1978).

Since then, quasispecies theory has been recognized as a subset of theoretical population genetics (Wilke 2005; Takeuchi and Hogeweg 2007). Recently, in the series of papers (Cerf 2015a, b; Dalmau 2015; Cerf and Dalmau 2016), it has been rigorously shown that the Wright–Fisher and Moran models for multi-loci mutation–selection converge to the single-peak fitness landscape quasispecies model, in the appropriate
limit of infinite populations. Moreover, due to its capability to accommodate high mutation rates, it has been widely applied to model the evolution of viruses with RNA genomes (Eigen 1993).

Inspired by the stochastic quasispecies model of Demetrius et al. (1985) and based on branching process techniques, Antoneli et al. (2013a, b) proposed a mathematical model aimed at understanding the basic mechanisms and phenomena of the evolution of highly mutating viral populations replicating in a single host organism, called **phenotypic (quasispecies) model**. It is denominated “phenotypic” due to the fact that it only comprises probabilities associated with the occurrence of deleterious, beneficial and neutral effects that operate directly on the replicative capability of viral particles, without any explicit reference to their genome. In Dalmau (2016), Dalmau introduced another generalization of the stochastic quasispecies model also based on multitype branching processes but retaining the genotypic character of Demetrius et al. (1985).

The phenotypic model (Antoneli et al. 2013a, b) is defined through a probability-generating function which formally determines the transition structure of the process. The matrix of first moments of the branching process, or simply the **mean matrix**, defines a deterministic linear system which describes the time evolution of conditional expectations, a “mean field model” for the actual stochastic process which is equivalent to the **Eigen’s selection equation** (Demetrius et al. 1985). The deterministic mean field model has been studied by several researches, but without the connection to a stochastic branching process, see for instance (Bergstrom et al. 1999; Manrubia et al. 2003; Cuesta 2011).

As shown in Antoneli et al. (2013a, b), the phenotypic model is fully specified by three fundamental parameters: the probabilities of occurrence of deleterious and beneficial effects \( d \) and \( b \) – the probability of occurrence of neutral effects is fixed by the complementary relation \( c = 1 - d - b \) – and the maximum replicative capability \( R \). By an exhaustive analysis of this “parameter space,” we were able to depict a fairly detailed portrait of all possible behaviors of the model. In Antoneli et al. (2013a), we carry out a thorough analysis of mean matrix, assuming that beneficial effects are absent and were able to show that the phenotypic model is “exactly solvable,” in the sense that the spectral problem for the mean matrix has an explicit solution. In Antoneli et al. (2013b), we employ **spectral perturbation theory** in order to treat the general case of small beneficial effects. This approach has provided a complete description of the generic behavior of the model.

In the present paper, we further address the biological implications of modeling RNA virus populations in terms of the phenotypic model. We achieved this goal by a combination of computational simulations and basic results of the theory of multitype branching processes as used in Antoneli et al. (2013a, b). In order to perform the simulations, we have implemented the phenotypic model into a computational platform capable of running the simulation and presenting the results in graphical format in real time.

We start with the description of the computational platform (its interface, output and main simulation routine). Then we proceed to use some of the theoretical results of Antoneli et al. (2013a, b) to validate the program with several simulation experiments that can be read independently from each other and are used to evaluate distinct
features of the program. Finally, we perform two additional simulation experiments to address the main questions of this paper:

(1) What is the impact of fitness distributions on the evolution of the phenotypic model and how to measure it?

(2) Is there an extinction mechanism similar to the “mutational meltdown” in the phenotypic model?

Role of fitness distributions

The fitness distributions of the phenotypic model are discrete distributions forming location-scale families parameterized by the replicative classes that control the progeny sizes at each replication cycle. They can be seen as representing distinct “compartments” in the host which can be more favorable or pose restrictions to the viral replication process. For instance, some distributions have a positive influence on the replication, by enhancing the replication of particles in the higher replicative classes, while other distributions have an opposite effect. Examples of favorable compartments are: existence of sites associated with immune privilege, or with lower concentration of antiviral drugs, or allowing for cell to cell virus transmission. Unfavorable compartments are sites with high antiviral drug penetration, small number of target cells, or accessed by elements of host responses as antibodies, cytotoxic cells and others. In this sense, we may think of fitness distributions as an environmental component during viral evolution. We show that the impact of the fitness distributions on the branching process is subtle and cannot be detected by quantities that depend only on the first moments of the process. Nevertheless, we introduce a new quantity, called *populational variance*, that is capable to detect the influence of different fitness distributions and is analytically and computationally tractable.

Unifying principle for extinction

According to the phenotypic model, a virus population can become extinct or eradicated from the host by the fulfillment of a condition involving only the probability of occurrence of deleterious effects $d$ and the maximum replicative capability $R$. Even further, in the absence of beneficial or compensatory effects, the fate of the population is determined by the product $R(1−d)$. If it is greater than 1, the population will survive or if it is lesser than 1 the population will face extinction. Based on this result, we show that there is a correspondence between two well-known distinct mechanisms of extinction:

(1) *Lethal Mutagenesis* (Loeb et al. 1999; Bull et al. 2007, 2008). The process of extinction of the viral population due to the increment of the deleterious rate.

(2) *Mutational Meltdown* (Lynch et al. 1993; Lynch and Gabriel 1990). The process of extinction of the viral population through the step-wise loss of the fittest replicative classes due to random drift associated with the finite population size effect.

The correspondence between the two mechanisms reinforces the view that both are “two sides of the same coin” (Matuszewski et al. 2017). We propose here an *unifying principle for the extinction of a virus population*: This principle is based on a mathematical model containing probabilities of neutral and deleterious effects and the average growth rate or average maximum fitness which is equivalent (under appropriate interpretation) to the extinction threshold of a branching process given by the malthusian parameter. In the course of the proof (see Sect. 3.3), we consider another parameter present in our phenotypic model, called the *carrying capacity*. Initially,
it was introduced as a convenient step for the computational implementation of the model, i.e., to prevent the population to grow boundlessly. Nevertheless, it can be seen as a genuine parameter of the model, which controls the intensity of the random drift. Because of this, we may consider our model as a self-regulated branching process, instead of a “pure” branching process. Furthermore, we observe that, even though the extinction mechanisms have the same mathematical “origin,” the processes leading to the actual extinction of the viral population may display distinct “signatures.”

Structure of the paper The paper is structured as follows. In Sect. 2, we introduce the computational platform for the simulation of the phenotypic model. In Sect. 3, we perform several simulation experiments to validate the program by comparing its output with the theoretical results from Antoneli et al. (2013a, b). We end this section with the presentation of the new results on the role of the fitness distributions and the mutational meltdown. The validation subsections and the two subsections on new results depend only on Sect. 2 and so can be read independently from each other. The paper ends with a conclusion section. There are 5 appendices. Appendices A, B and C provide some background on branching process theory and theoretical results about the phenotypic model for the reader’s convenience. Appendices D and E provide some details about the implementation of the computational platform introduced in the paper.

2 Software Description

In this section, a computational platform is introduced for the simulation of the phenotypic model of Antoneli et al. (2013a, b).

2.1 The ENVELOPE Program

The ENVELOPE (EvolutioN of Virus populations modELed by stOchastic ProcEss) program is a cross-platform application developed to simulate the phenotypic model of Antoneli et al. (2013a, b). The software contains a graphical interface to input data, visualize graphics in real time, and export the output data to CSV format, which can be used with a wide range of statistical analysis tools. It was written in C++ programming language using the Qt framework to design the graphical user interface. It was exhaustively tested on Linux operating systems.

The main window of the program has several tabs with the first called “Data Input” where the user can set the values of several parameters that completely specify the model, as follows (see Fig. 1).

– Total probability ($u$): the probability that a progeny particle will undergo some fitness effect. It should be a number between 0 and 1. The effect of this probability is to renormalize the other probabilities ($p \mapsto u \cdot p$), and its default value is $u = 1$ (no renormalization).

– Beneficial probability ($b$): the probability of occurrence of a beneficial effect. It should be a number between 0 and 1.
- **Deleterious probability** \((d)\): the probability of occurrence of deleterious effect. It should be a number between 0 and 1. The complementary probability \(c = 1 - b - d\) is the probability of occurrence of neutral effect. If \(b + d > 1\), then \(c\) is set to 0 and \(d = 1 - b\).
- **Replicative classes** \((R)\): the number of nonzero replicative classes, and hence there are \(R + 1\) replicative classes (maximum replicative capability).
- **Max population size** \((K)\): the maximum population size (carrying capacity).
- **Max generation time** \((N)\): the total number of generations to be simulated. Each generation corresponds to a replication cycle.
- **Multi-core processor**: controls the recruitment of processors by the program.
- **Initial population**: the number of particles in each replicative class that will initiate the process.
- **Distribution**: location-scale family of fitness distributions (see Table 1).

The remaining tabs (“Progeny,” “Class Distribution,” “Average,” “Diversity,” “Entropy,” “Variance”) display graphics of the above quantities in real time as the simulation proceeds. The tab “Data Output” displays a table with all the data generated during the simulation. These data can be saved to a file (button “Save to File”) or copied to the memory (button “Copy to Memory”), and then it can be directly pasted into a spreadsheet.

The button “Process” starts the simulation, the button “Finish” ends the simulation at any time, and the button “Exit” closes the program. If the total number of particles in a generation is equal to zero, it is assumed that the population has become extinct,
and hence the simulation stops. The button “Video” pauses the simulation, without ending the simulation, and allows the user to change the above parameter settings and continue the simulation with the new setting. This feature is used to emulate the changes in the environment – the host organism – where the reproduction process takes place.

The evolution of the population can be measured through a few simple quantities that vary as a function of the generation number \( n \geq 0 \). Let \( \mathbf{Z}_n = (Z_0^n, \ldots, Z_R^n) \) denote the vector whose component \( Z_r^n \) is the number of particles in the \( r \)th replicative class at generation \( n \).

- **Progeny size**: total number of particles \( |\mathbf{Z}_n| = \sum_r Z_r^n \) at generation \( n \).
- **Relative growth rate**: the relative growth rate at generation \( n \) given by (for \( n > 1 \))

\[
\mu(n) = \frac{|\mathbf{Z}_n|}{|\mathbf{Z}_{n-1}|}
\]

It is a multidimensional version of the Lotka–Nagaev estimator (Lotka 1939; Nagaev 1967), which gives an empirical estimator of the malthusian parameter.

- **Asymptotic distribution of classes**: the proportion of particles in the \( r \)th replicative class at generation \( n \) given by

\[
u_r(n) = \frac{Z_r^n}{|\mathbf{Z}_n|}
\]

The vector \( \mathbf{u}(n) = (u_0(n), \ldots, u_R(n)) \) is called asymptotic distribution of classes (or simply the class distribution).

- **Average reproduction rate**: the average reproduction rate (mean of the class distribution) at generation \( n \) given by

\[
\langle \rho(n) \rangle = \sum_{r=0}^{R} r u_r(n)
\]

It can be shown that the average reproduction rate equals to the relative growth rate:

\[
\langle \rho(n) \rangle = \mu(n) \quad \text{for all} \quad n > 1
\]

(see “Appendix A” for details).
– Phenotypic diversity: the variance (or standard deviation) of the class distribution at generation $n$ given by

$$
\sigma^2_{\rho}(n) = \sum_{r=0}^{R} r^2 u_r(n) - \langle \rho(n) \rangle^2
$$

– Phenotypic entropy: the informational or Shannon entropy of the class distribution at generation $n$ given by

$$
h_{\rho}(n) = - \sum_{r=0}^{R} u_r(n) \ln u_r(n)
$$

Here, we use the convention “$0 \ln 0 \equiv 0$.” This quantity behaves very much like the phenotypic diversity.

– Normalized populational variance: the normalized populational variance at generation $n$ given by

$$
\phi(n) = \sigma^2(n) - \sigma^2_{\rho}(n)
$$

where $\sigma^2$ is the empirical estimator of the variance corresponding to the malthusian parameter $\mu(n)$ (see “Appendix A” for details).

Strictly speaking, a surviving population described by branching process which does not becomes extinct grows indefinitely, at an exponential rate proportional to $\mu^n$. Hence, in order to simulate a branching process it is necessary to impose a cutoff on the progeny size, otherwise it would blow up the memory of the computer. This cutoff is done by setting the maximum population size $K$ which controls how much the population can grow unconstrained, acting in a similar fashion as the carrying capacity of the logistic growth (Campbell 2003; Lambert 2005). If the total number of particles that comprises the current generation is greater than the maximum population size $N$, a random sampling procedure is performed to choose $N$ particles to be used as parental particles for the next generation. In particular, the progeny growth curve resembles a “Logistic Growth Curve” (see Fig. 2).

There are also some other additional settings that alter the way the program behaves. “Produce zero class particles” allows to set whether the particles of replicative capability $r = 0$ will be considered in the calculations or not. “Previous last generation/Do not preserve last generation” allows to choose whether the particles in previous generation will be carried over to current generation. This was included in order to account for the possibility of a replication strategy that does not implement the disassemble of the parental particle. In most cases, the replication strategy used by RNA viruses implements the disassemble of the virus particle during the replication. Retroviruses replication process is performed by the reverse transcriptase enzyme. The process of reverse transcription involves the synthesis of complementary DNA from the single-stranded RNA followed by the degradation of the intermediate RNA–DNA hybrid form. The preservation of the parental generation in the model of viral evolution can
allow one or more particles to be preserved during several generations, in contrast with the above-mentioned replication strategies of the RNA viruses. The main routine of the program is given by the pseudo-code in “Appendix E.”

Finally, in order to discuss the simulations for the case when \( b = 0 \) it is useful to introduce some conventions. The \textit{instantaneous maximum replicative capability (at generation \( n \))}, defined by \( r^* (n) = \max \{ r : Z^r_n \neq 0 \} \), where \( Z_n = (Z^0_n, \ldots, Z^R_n) \) is the vector whose component \( Z^r_n \) is the number of particles in the \( r \)th replicative class at generation \( n \geq 0 \). If the initial population \( Z_0 = (Z^0_0, \ldots, Z^R_0) \) has \( r^* (0) < R \), then all the quantities that depend on \( R \) can must be calculated with \( r^*_n \) in the place of \( R \), at the generation \( n \). Note that if \( b = 0 \) then, for all purposes, \( r^*_n = r^*_0 \) acts as the maximum replicative capability. Even when \( b \neq 0 \), the parameter \( r^*_n \) acts as an “instantaneous” maximum replicative capability, which changes only when a particle in the highest replicative class \( r^*_n \) produces a progeny particle in the next replicative class, namely \( r^*_n (n + 1) = r^*_n (n) + 1 \), that is retained in the population.

3 Simulation Experiments

In this section, we use some of the theoretical results from Antoneli et al. (2013a, b) (see also “Appendix B”) to validate the \textsc{envelope} program at several levels of refinement. The validation is subdivided into several parts corresponding to distinct features of the model that are classified according to the possible regimes and phases of the time evolution of a multitype branching process. The subitems of the validation subsection can be read independently from each other. We also present new consequences of the
combination of simulations with theoretical analysis and provide new perspectives on the role of fitness distributions and the variance of the branching process and on the mechanism of extinction, allowing us to propose a unifying principle underlying for the extinction of a virus population.

3.1 Validation of the ENVELOPE Program

3.1.1 Transient Phase and Recovery Time

A heterogeneous population replicating in a constant environment typically undergoes an initial period of high stochastic fluctuations in the relative frequency of each variant, until it reaches a stationary regime where the relative frequencies become constant. This initial period, called transient phase, is marked by the beginning of the viral infection, after the bottleneck event when one or more particles are transmitted to a host organism and initiates the process of (re)establishment of the viral population in the new host. The transient phase comprises the acute infection phase (Fiebig et al. 2003; McMichael et al. 2010), which is characterized by an initial exponential growth of the population, the attainment of the viremia peak, followed by a slower decrease toward a stabilization of the population size (see Fig. 2).

In the phenotypic model, the transient regime corresponds to the beginning of the time evolution of the process. It is characterized, as noted before, by an instability of the relative frequencies of the replicative classes, an exponential growth of the progeny size, a decrease in the average reproduction rate, and an increase in both the phenotypic diversity and the phenotypic entropy.

The expected time (as function of the number of generations) of the relaxation toward an equilibrium after the bottleneck event, called recovery time (see the initial segment of the time series in Fig. 3).

Let us assume, as usual, that the beneficial probability $b \approx 0$ and the founding population have $r_*(0) < R$. Then one observes that the progeny size, the average reproduction rate, the phenotypic diversity and phenotypic diversity display a time series with several plateaus. The “length” of each plateau is the number of generations that the population remains with the same value of $r_*(n)$ and the higher the plateau, the longer, on average, is its length. The presence of a jump indicates that a progeny particle form a parental particle in the replicative class $r_*(n)$ has undergone a beneficial effect, that is, the active maximum replicative capability increases by 1 unit: $r_*(n + 1) = r_*(n) + 1$. The occurrence of jumps can go on until $r_*(n) = R$. Therefore, the “length” of each plateau represents the time, in number of generations, required for a beneficial effect to occur on a particle at the highest replicative class and be retained in the population. The probability $P(\text{jump in } r_*(n))$ of occurrence of a jump event, when $r_*(n) < R$, may be estimated using Eq. (22) of “Appendix B” as

$$P(\text{jump in } r_*(n)) \approx b u_{r_*(n)} \approx b \left(1 - d\right)^{r_*(n)}$$

where $u_{r_*(n)}$ is the proportion of particles in the $r_*(n)$th replicative class at generation $n$, which is the instantaneous maximum replicative capability at time $n$. Notice that, as
Average reproduction rate = relative growth rate. The “jumps” associated with the recovery time have heights about 0.5. Parameter values: $b = 0.000001; d = 0.50; Z_0^2 = 1; R = 10; N = 4000; K = 10^6$; fitness distribution: Delta.

$r_*(n)$ increases, $u_{r_*(n)}$ decreases monotonically, and therefore, $P(\text{jump in } r_*(n)) \to 0$ when $r_*(n) \to \infty$. This result highlights the asymmetry between the contributions of the beneficial probability versus the deleterious probability to the recovery time.

The “height” of a jump in the average reproduction rate time series is independent of the plateau where the jump occurs. In order to estimate the “height,” consider two consecutive levels on the time series of the average reproduction rate, the first “height” $\mu(n_1)$ measured at generation $n_1$ and the second “height” $\mu(n_2)$ measured at generation $n_2$, with $n_1 < n_2$ not necessarily consecutive, such that $r_*(n_2) = r_*(n_1) + 1$ and $\mu$ is approximately constant around $n_1$ and $n_2$. Thus, the difference $\mu(n_2) - \mu(n_1)$ gives an estimate of the height of the jump between two consecutive plateaus. When $b \approx 0$, Eq. (20) of “Appendix B” implies that $\mu(n) \approx r_*(n)(1 - d)$, and hence

$$\mu(n_2) - \mu(n_1) \approx (r_*(n_2) - r_*(n_1))(1 - d) \approx 1 - d.$$ 

For instance, in Fig. 3 it can be readily seen that the height of the jumps is about 0.5 and, in fact, $d = 0.50, b = 0.000001$, and hence $1 - d = 0.4999999$.

### 3.1.2 Stationary Regime

The advanced stage of the infection, also called chronic infection phase (Fiebig et al. 2003; McMichael et al. 2010), is comprised by the stationary regime where the viral population has recovered its phenotypic (and genotypic) diversity and becomes better adapted to the new host environment by exhibiting rather stable relative frequencies of almost all variants.
In the phenotypic model, the stationary regime corresponds to the asymptotic behavior of a super-critical branching process ($\mu > 1$). But, as mentioned before, a surviving population described by super-critical branching process is never stationary (in the strict sense), and therefore this correspondence is not straightforward.

The normalized process $W_n = Z_n / \mu^n$ is stationary and, when $n \to \infty$, the random variable $Z_r^n / |Z_n|$ converges to the asymptotic relative frequency $u_r$ of $r$th replicative class. Consequently, the average reproduction rate $\langle \rho(n) \rangle = \mu(n)$, the phenotypic diversity $\sigma^2\rho(n)$ and the phenotypic entropy $h_\rho(n)$ remain essentially constant in time. Moreover, the maximum population size cutoff $K$ ensures that the total progeny size remains constant in time with expected value $\langle |Z_n| \rangle \approx \mu(n) K$.

During the stationary regime, the stability of the relative frequency of each class is maintained by a steady “flow of particles” from a replicative class to its adjacent classes, due to the deleterious probability $d$ and the beneficial probability $b$. The probability $c$ contributes maintenance of a constant proportion of particles in each replicative class. When the beneficial probability $b \neq 0$ the asymptotic distribution of classes $u_r$ is independent of the configuration of the founding population and, when $n$ is large enough, $r^*(n) = R$.

More importantly, when $b \approx 0$, the replicative classes that are most representative in the population are the classes near the mode of the distribution of classes $u_r$, also known as “most probable replicative capability.” The mode of $u_r = \binom{r; R}{1 - d}$ is given by $m(u_r) = \lfloor (R + 1)(1 - d) \rfloor$, except when $(R + 1)(1 - d)$ happens to be an integer, then the two replicative classes corresponding to $(R + 1)(1 - d) - 1$ and $(R + 1)(1 - d)$ are equally “most probable” (see Feller 1968; here, $\lfloor x \rfloor$ denotes the greatest integer less than $x$). When $(1 - d) \approx 1/2$, the mode is close to the average reproduction rate $\mu(n) = \langle \rho(n) \rangle$ (see Fig. 4).

### 3.1.3 Threshold of Extinction

The threshold of extinction takes place when the deleterious rate is sufficiently high that it prevents the viral population of reaching the stationary regime but not high enough to induce the extinction of the population in the short run. Therefore, any small increase in the deleterious rate can push the population toward extinction, while any small decrement can allow the population to reach the stationary regime.

In the phenotypic model, the threshold of extinction corresponds to a critical branching process ($\mu = 1$) and is characterized by instability of the relative frequencies of the replicative classes, the average replicative rate and the phenotypic diversity. The instability observed represents the impossibility of the viral population to preserve, due to the deleterious effects, particles with high replicative capability. The occurrence of an eventual extinction of the population is almost certain, although the time of occurrence of the extinction may be arbitrarily long if the initial population is sufficiently large. In other words, the threshold of extinction looks like an infinite transient phase and is the borderline between the stationary regime, where the transient phase ends at an stationary equilibrium, and the extinction in finite time.

Setting the parameters of the phenotypic model in order to obtain a critical branching process is a matter of “fine tuning,” since it requires that the probabilities $d$, $b$ and ...
the maximum replicative capability $R$ satisfy the algebraic equation $\mu(b, d; R) = 1$ – which is a non-generic condition (see Fig. 5).

When $b = 0$, the critical deleterious probability is $d_c = 1 - 1/R$, for each fixed $R$. When $b \neq 0$ one may consider, for fixed $R$, the corresponding critical probabil-
Table 2. Critical deleterious probabilities $d_c(0)$ and $d_c(b^*)$.

| $R$ | $d_c(0)$ | $d_c(b^*)$ | $\tilde{d}_c(b^*)$ | $|d_c(b^*) - \tilde{d}_c(b^*)|$ |
|-----|----------|------------|---------------------|---------------------------------|
| 2   | 0.50     | 0.707      | 0.750               | 0.043                           |
| 3   | 0.66     | 0.895      | 0.933               | 0.038                           |
| 4   | 0.75     | 0.951      | 0.975               | 0.024                           |
| 5   | 0.80     | 0.972      | 0.988               | 0.016                           |
| 6   | 0.83     | 0.982      | 0.993               | 0.011                           |

The real values of $d_c(b^*)$ were obtained by numerical computation using the mean matrix, and the values denoted by $\tilde{d}_c(b^*)$ were obtained using Eq. (1).

The critical deleterious probability $d_c$ as given implicitly by the equation $\mu(b, d_c(b), R) = 1$ and the condition $\mu(0, d_c(0), R) = 1$, with $d_c(0) = 1 - 1/R$ (see Fig. 10 of “Appendix B”). Since $b$ and $d$ are constrained to satisfy $b + d \leq 1$, there is a maximum value of $b$ such that $b + d_c(b) = 1$, for each fixed $R$. Denote this maximum value by $b^*(R)$. The number $b^*(R)$ is the maximum beneficial probability such that the phenotypic model has three distinct regimes. In other words, if $b > b^*(R)$ then $d < d_c(b)$ and the process never becomes extinct. In the parameter space of the phenotypic model, the critical probabilities $d_c(b^*)$ at $b^*$ are given as the intersection of the boundary line $b + d = 1$ with the critical curves $\mu(d, b, R) = 1$ for each fixed $R$.

Using the expressions for the malthusian parameter obtained in Antoneli et al. (2013b), it is easy to show that the following approximations hold (when $R \to \infty$)

\[
b^*(R) \approx \frac{1}{R} - \frac{1}{R} \frac{(R-1)^2}{1 + (R-1)^2}
\]

\[
d_c(b^*(R)) \approx d_c(0) + \frac{1}{R} \frac{(R-1)^2}{1 + (R-1)^2}.
\]

(1)

Here, one uses that $b^* + d_c(b^*) = 1$ and $d_c(0) = 1 - 1/R$. Comparison of critical deleterious probability given by Eq. (1) with the correct values obtained by numerical computation using the mean matrix, shown in Table 2, indicates that the asymptotic expressions converge to the real values when $R \to \infty$.

3.1.4 Extinction by Lethal Mutagenesis

The process of extinction of the viral population induced by increase in the deleterious rate is called lethal mutagenesis (Bull et al. 2007). In the phenotypic model, the lethal mutagenesis corresponds to a sub-critical branching process ($\mu < 1$). It is characterized by continuous decrease in the average replicative rate and by increase in the phenotypic diversity followed by a sudden decrease in the subsequent generations.

The progeny size and the phenotypic diversity increase during the first generations because the founding population still has reasonable replicative capability. However, increasing the size of the founding population does not prevent extinction, it only increases the time required for the extinction to occur. Increasing the deleterious probability $d$ decreases the time required for extinction, and increasing beneficial probability $b$ can prevent extinction.
Fig. 6 Lethal mutagenesis and the path to extinction. Parameter values: $b = 0; d = 0.501; R = 2; N = 2500; K = 10^6; Z_0 = (1000, 2000, 1000);$ fitness distribution: Delta

Note that when $b = 0$, the population cannot achieve a replicative capability higher than the one present in the founding population. In this case, a population transmitted to a new host organism via a bottleneck event will have maximum replicative capability less or equal to the maximum replicative capability of the original population.

Interesting enough, there is a signature of the extinction process which may be directly observed in the behavior of the average reproduction rate curve $\mu(n)$. It is marked by an explosive growth in the variation of $\mu(n)$ as $n$ approaches the extinction time $n^*$ (see Fig. 6).

The phenomenon of explosive growth near the extinction event may be detected by the oscillation of $\mu(n)$ in an interval ending at the last nonzero generation:

$$\text{osc}(\mu) = \max_{n < n^*} \mu(n) - \min_{n < n^*} \mu(n).$$

Even when the process is slightly super-critical, it is expected that the oscillation of $\mu(n)$ remains very small, with $\text{osc}(\mu) \sim 10^{-3}$ for all $n$. On the other hand, when a slightly sub-critical process is approaching the extinction time $n^*$ one typically observes $\text{osc}(\mu) \sim 10^{-1}$.

The expected time to extinction $\langle T_{\text{ext}} \rangle$ of a branching process was determined in Jagers et al. (2007): if $\mu \leq 1$ then

$$\langle T_{\text{ext}} \rangle = \frac{\ln Z_0^{r\ast} + \kappa}{-\ln \mu},$$

where $\kappa > 0$ depends only on the parameters of the model (not on the initial population). It is easy to show that at the critical value of the malthusian parameter
(\mu = 1) equilibrium is never reached. A scaling exponent characterizing the behavior of expected time to extinction in a neighborhood of the critical value of the malthusian parameter can be obtained by considering the first-order expansion of \langle T_{\text{ext}} \rangle about 1:

\langle T_{\text{ext}} \rangle \approx |\mu - 1|^{-1}.

When \( b = 0 \) one may write \langle T_{\text{ext}} \rangle as a function of the deleterious probability and the critical deleterious probability \( d_c = 1 - 1/R \) as

\langle T_{\text{ext}} \rangle \approx \frac{1}{R} |d - d_c|^{-1}

since |\mu - 1| = R|d - d_c|. This “scaling law” is formally identical to the one obtained in Gupta and Dixit (2015) for the error threshold of the deterministic quasispecies model as a function of the mutation rate.

### 3.2 Populational Variance and the Role of Fitness Distributions

All properties of the phenotypic model that have been discussed so far are related to the mean matrix of the model, that is, they depend only on the first moments of the branching process and may be called “first-order properties.” In particular, they are independent of the choice of the family of fitness distributions. If we want to see how the fitness distributions influence the evolution of the population, we must to look at a “second-order property,” which is expected to depend on the second moments of the fitness distributions (see Table 1).

The simplest property of second order is given by the population variance \( \sigma^2 \) associated with the malthusian parameter \( \mu \) (namely the relative growth rate). Furthermore, the difference between the populational variance and the (squared) phenotypic diversity, called normalized populational variance and denoted by \( \phi \) is a very interesting quantity to be measured, since it satisfies

\begin{equation}
\phi = \sigma^2 - \sigma_{\rho}^2 = \sum_{r=0}^{R} \sigma_r^2 u_r
\end{equation}
In other words, $\phi$ is a weighted average of the variances $\sigma^2_r$ of the fitness distributions. See “Appendix A” for the precise definition of $\sigma^2$ and the proof of the second equality in Eq. (2).

Given a location-scale family of fitness distributions $t_r$ such that $\sigma^2_r$ is at most a quadratic polynomial on $r$, Eq. (2) allows one to write the corresponding normalized population variance $\phi$ in terms of the average reproduction rate $\langle \rho \rangle$ and the phenotypic diversity $\sigma^2_\rho$. Hence, $\phi$ can be exactly computed for all location-scale families of distributions used in the ENVELOPE program (see Table 3).

It is important to stress that unlike the malthusian parameter, the normalized population variance does depend on the choice of the family of fitness distributions. Recall that the malthusian parameter depends only on the mean matrix, which depends on the fitness distributions $t_r$ only through its expectation values. Since we have imposed the same normalization condition that the expectation value of $t_r$ is $r$ for all families of fitness distributions, it follows that the mean matrix, and hence the malthusian parameter, does not depend on the family of fitness distributions. On the other hand, the variances of different families of fitness distributions are not necessarily the same. For instance, if $t_r$ is the family of Poisson distributions then $\sigma^2_r = r$ and thus $\phi = \mu$.

Assume that $b = 0$ (then $c = 1 - d$). From the expression of the asymptotic distribution of classes (21), one obtains: $\langle \rho \rangle = \mu = R(1 - d)$ and $\sigma^2_\rho = Rd(1 - d)$. Moreover, when $b \neq 0$ is sufficiently small, formula (22) ensures that $\langle \rho \rangle$ and $\sigma^2_\rho$ are approximated by the corresponding values for $b = 0$ and the same holds for $\phi$.

For instance, in Fig. 7 we show the graph of the normalized population variance $\phi(n)$, at generation $n$, from a simulation in which we switched among the four families of fitness distributions with finite variance using the “Video” function of the

---

**Fig. 7** Normalized population variance $\phi$, with $\langle \rho \rangle = 1$ and $\sigma^2_\rho = 0.5$. Parameter values: $b = 0; d = 0.50; R = 2; N = 2000; K = 10^6; Z^2_0 = 10,000$; fitness distributions: Delta ($\phi = 0$), Poisson ($\phi = 1$), Geometric ($\phi = 2.5$), Binomial ($\phi = 0.5$)
ENVELOPE program to pause the simulation and change the type of fitness distribution.

Finally, it is worth to remark that the impact of the power law family of fitness distribution on the evolution of the population is very distinct from the other families, because, unlike the other fitness distributions, it has infinite variance. One of the consequences of this property is the appearance of intense bursts of progeny production clearly seen on the times series of progeny size and the average reproduction rate (see Fig. 8). The instability caused by unbounded fluctuations coupled with the finite population size effect (even for large \( K \)) is responsible for the generation of a train of sparse and intense bursts of progeny production. On the other hand, this instability coupled with finiteness effect may also provoke sudden drops on the progeny size driving the population to a premature extinction, even if the malthusian parameter is above 1. Because of these extreme phenomena, one would be led to believe that the phenotypic model with the power law family of fitness distributions is an exception to the general result: Any property derived from the mean matrix is independent of the fitness distribution. It is not the case. In fact, if one considers the time-average of any quantity that is time-dependent over a time interval \([n_0, N]\) during the stationary regime, let’s say

\[
\bar{\mu}(N) = \frac{1}{N} \sum_{n=n_0}^{N} \mu(n),
\]

then it is expected that \(\bar{\mu}(N)\) becomes very close to the asymptotic value of the relative growth rate \(\mu\) when \(N\) is sufficiently large.

For instance, in Fig. 8 the time-average of the progeny size \(\bar{\mu}(N)\) is around 14,000, while the expected progeny size for the model is \(\mu K = R(1-d)K = 2 \times 0.7 \times 10^4 = 1.4 \times 10^4 = 14,000\), in full agreement with the general theory.

### 3.3 Finite Population Size and Mutational Meltdown

Recently, Matuszewski et al. (2017) reviewed the literature about theories and models describing the extinction of populations owing to the excessive accumulation of deleterious mutations or effects and distinguished two apparently distinct lines of research, represented by the lethal mutagenesis models (Bull et al. 2007) and the mutational meltdown models (Lynch and Gabriel 1990) which, nonetheless, display a considerable amount of similarity.

Indeed, as shown in Bull et al. (2007), Antoneli et al. (2013a, b), lethal mutagenesis is independent of population size, and hence it is fundamentally a deterministic process that operates even on very large populations. Although the outcome of lethal mutagenesis is deterministic, other aspects of the population dynamics (such as extinction time, individual trajectories of progeny size) are not. On the other hand, the mutational meltdown generally works within the context of “small” population sizes in which stochastic effects caused by random drift play an important role.

We believe that the approach presented here may help shed some light on this issue. There is one ingredient in the mutational meltdown theory that is absent in
the lethal mutagenesis theory: the carrying capacity. This is true even for models with finite population, such as (Demetrius et al. 1985) and the phenotypic model, in their theoretical formulations as branching process. However, as seen before, the computational implementation of the phenotypic model required the introduction a cutoff $K$ in order to bound the growth of the population. If the cutoff is taken as basic constituent of the phenotypic model, and not merely a convenient device, then it can play a role similar to a carrying capacity and the model may no longer be considered a “pure” branching process, but a self-regulating branching process (Mode and Sleeman 2012; Mode et al. 2013).

In a self-regulating branching process, not all the offsprings produced in a given generation will produce offspring in the next generation, and hence, it is necessary to introduce a survival probability distribution $S(n|T_n)$, to stochastically regulate the survival of offspring at any generation $n$ as a function of the total population size $T_n = |Z_n|$. The motivation behind this definition is as follows: If the population size at a generation $n$ exceeds the carrying capacity of the environment, then, due to competition for resources, it is less likely that an offspring produced in that generation will survive to produce offspring at generation $n + 1$.

Let $S(n|T_n)$ denote the conditional probability that any offspring produced at generation $n$ survives to produce offspring at generation $n + 1$, given that the population has $T_n$ individuals at generation $n$. If we define the conditional probability $S$ as

$$S(n|T_n) = \begin{cases} \frac{K}{T_n} & \text{if } T_n > K \\ 1 & \text{if } T_n \leq K \end{cases}$$
then the phenotypic model becomes a self-regulating process with carrying capacity $K$. Moreover, when $K \to \infty$ the self-regulating process reduces to a “pure” branching process.

If $K$ is not large enough, then a kind of random drift effect due to finite population size may take place, which happens when the fittest replicative classes are lost by pure chance, since its frequency is typically very low. (They are the lesser represented replicative class in the population.) If the loss of the fittest replicative class occurs a sufficient number of times, then the population will undergo extinction. Note that this may happen even when the process is super-critical, namely it is far from the extinction threshold. This is not a contradiction with the definition of extinction probability, since a super-critical process still has a positive probability to become extinct (see “Appendix A”).

Now suppose that $b = 0$, the initial population has active maximum replicative capability $r_n(0)$ and the carrying capacity $K$ is sufficiently small. (We shall give an estimate of $K$ in a moment.) Then, as mentioned before, the value $r_\ast = r_n(0)$ acts as the maximum replicative capability for that population. Moreover, if the highest replicative class $r_\ast$ is lost by chance, that is, if $r_\ast(n + 1) = r_\ast(n) - 1$, then it cannot be recovered anymore, and hence, from that time on the maximum replicative capability for that population has dropped by 1 unit. This may be seen as a manifestation of the “Muller’s ratchet,” since the population has accumulated a deleterious effect in an irreversible manner.

For sake of concreteness, let us assume that $r_\ast = R$ and $d$ are such that $(R - 1)(1 - d) < 1$, but $R(1 - d) > 1$. Then, at the beginning of the process, the malthusian parameter is $\mu = R(1 - d) > 1$ and the process is super-critical. However, if at some generation $n$, the $R$th replicative class is lost by chance, then $R$ drops by 1 and $\mu = (R - 1)(1 - d) < 1$, so the process becomes sub-critical and the population becomes extinct very quickly. In this case, the frequency of the $R$th replicative class is $(1 - d)^R$ and fraction of particles that are purged, at each generation, is $R(1 - d) - 1$; hence, the fraction of particles that are left in the $R$th replicative class, at each generation, is $v_R = 2(1 - d)^R - R(1 - d)^{R+1}$. If $K \approx 1/v_R$, then there will be, on average, 1 particle of class $R$ per generation – it is very unlikely that this replicative class will be retained for a long period of time. Therefore, in order to avoid the random drift effect $K$ should be at least of the order of $10 \times R(1 - d)/v_R$, or higher. At each “click of the ratchet,” the fittest replicative class is lost and there is a drop in the malthusian parameter by $(1 - d)$, until $r_\ast(1 - d)$ becomes less than 1, where $r_\ast$ is the maximum replicative capability at the current generation. This drop occurs in the phenotypic diversity and the phenotypic entropy, as well (see Fig. 9).

If one writes the usual condition for occurrence of extinction $r_\ast(1 - d) < 1$ as

$$(1 - d) < 1/r_\ast$$

then this is an exact phenotypic analog of the mutational meltdown extinction criterion. Indeed, $r_\ast$ is the phenotypic analog of absolute growth rate of the population at time $n$ and $c = (1 - d)$, the probability of occurrence of a neutral fitness effect per individual particle, is the phenotypic analog of the mean viability [compare with the equations in Lynch et al. (1993), Lynch and Gabriel (1990)].
Fig. 9  Extinction by mutational meltdown. Phenotypic Entropy time series. At the beginning, the branching process is super-critical with $\mu = 3.41$. Parameter values: $b = 0; d = 0.659; R = 10; N = 10,000; K = 2000; Z_{10}^0 = 6000; \text{fitness distributions: Delta}$

4 Conclusion and Outlook

In this paper, we have exhaustively explored a model for the evolution of RNA virus, which was formulated as a multivariate branching process, called *phenotypic model*. The theory of branching processes provides a suitable framework endowed with concepts and analytic tools allowing for the investigation of evolutionary aspects of RNA viruses propagating along different adaptive landscapes.

One of the greatest virtues of the phenotypic model is its simplicity. Since the model has essentially only 3 parameters, it is possible to analytically compute the spectrum of its mean matrix and, applying the classification of multitype branching processes, obtain a complete qualitative description of its “generic behaviors,” that is, the most likely outcomes of the model’s asymptotic dynamics.

The maximum replication capacity $R$ and the probabilities of occurrence of deleterious effects $d$ entirely determine whether a viral population becomes extinct infinite time or not. On the other hand, the third parameter, the probability of occurrence of beneficial effects $b$, plays a distinct role from the other two probabilities, functioning as a threshold parameter which determines whether the model posses the three typical regimes of a branching process or just one regime (super-critical).

The model provides several statistical measures, such as average growth rate, phenotypic diversity, phenotypic entropy and population variance, that allows one to assess the stochastic dynamics of a viral population. The dynamics of the associated deterministic quasispecies model is given by a mean field limit where the mean matrix completely determines the dynamics (see “Appendix C”). Hence, it is possible to establish a relation between statistical measures mentioned above and the fundamental macroscopic...
parameters that characterize the evolutionary dynamics of a quasispecies. In particular, the models of Swetina and Schuster (1982), Schuster and Swetina (1988) could be used as a representation of the evolution of mean values obtained from the mean matrix of a branching process. By extending the scope of the model to age-dependent branching processes, (Athreya and Ney 1972) could allow the incorporation other statistical measures, such as evolutionary entropy (Dietz 2005; Demetrius 2013). This quantity could provide a more precise understanding of viral diversity given the fact that population sizes of viral population are finite.

Despite its conceptual appeal, the phenotypic model has some important drawbacks. The first limitation is the lack of feedback from the host organism on the virus population, since the probabilities of fitness effects are independent of time. This shortcoming is partially handled in the ENVELOPE program by the “Video” function, which allows one to pause the simulation and change the probabilities and emulate the host’s “response” against the virus. The second limitation is the lack of the phenotype-to-genotype map, i.e., the relationship between genotypic and phenotypic change. The motivation to use a phenotypic approach was to avoid the severe difficulties in modeling this kind of mapping (Alberch 1991; Fortuna et al. 2017).

Even though there is no phenotype-to-genotype map, it still is possible to draw some consequences about mutation rates from a purely phenotypic model (Antoneli et al. 2013a). For example, under the assumption that the mutation rate $U$ is sufficiently high (between 0.1 and 1), the probability that a spontaneous mutation produces a deleterious effect may be estimated as follows: If we assume that the number of mutations in a genome follows a Poisson distribution, then $d_{sp} \approx 1 - e^{-f_d}$, where $f_d$ is the probability that a spontaneous mutation has a deleterious effect (Kimura and Maruyama 1966). Values of $f_d$ have been measured in vitro for a few viruses and are shown in Table 4, along with the respective mutation rates $U$. Now, $d_{sp}$ provides a lower bound for the deleterious probability and since the value $d_{sp} \approx 1/2$ seems to be typical for RNA viruses, the interval $1/2 < d < d_c = 1 - 1/R$ is more likely to be the range of the parameter $d$. Moreover, it is easy to see that the phenotypic diversity and the phenotypic entropy are maximal when $d$ is near $1/2$, for any value of $R$ (Antoneli et al. 2013a). One could speculate that this is a universal property for RNA viruses that replicate under high mutational rates associated with a maximization principle that seeks to improve the chances of survival (Drake 2012).

The first main result of this paper concerns the role of the fitness distributions. The fitness distributions of the phenotypic model were motivated by the results and observations of Zhu et al. (2009) on the distribution of single-cell progeny sizes of RNA viruses. In Zhu et al. (2009), the authors demonstrated that even in a well-controlled experiment, using the same viral isolate, same infection parameters and clonally expanded target cells, progeny sizes can vary substantially. The variance on progeny sizes in such uniform environment indicates that RNA viruses replication bears in some way a portion of unpredictability. In this manner, it is impossible to know how many particles will be produced by a cell until the infection takes place and the progeny is released. Thus, fitness distributions provide a simple way to accommodate this unpredictability into each viral replication cycle. In fact, as shown here, some types of fitness distributions may have a substantial impact on the evolution of the viral population, most notably the power law. The extreme behavior produced by
the power law resembles that of the “viral load blips” frequently observed in HIV patients under highly active antiretroviral therapy (HAART) (Di Mascio et al. 2003; Nettles et al. 2005; Nettles and Kieffer 2006; Lee et al. 2006; Gallant 2007; Rong and Perelson 2009b, a), with undetectable or very low viral loads. This particular prediction of the model agrees with the assumption that these events are simply due to random fluctuation of the replication process, since after a “blip” the viral load quickly returns to its basal values.

The second main result of this paper is concerned with the mechanisms that drive a RNA virus population to extinction. As mentioned before, in the framework of multitype branching processes, these are essential to main attributes associated with this type of event: the probability of occurrence of deleterious effects $d$ and the maximum replicative capability $R$. If in addition to these two, one also considers the carrying capacity as a fundamental parameter of the phenotypic model then it is possible to show that the two principal mechanisms of extinction, lethal mutagenesis and mutational meltdown, are based on the same mathematical principle. Therefore, as far as the phenotypic model is concerned, this is a proof of the claim (Matuszewski et al. 2017) that these two mechanisms are “two sides of the same coin.”

Finally, for the sake of simplicity the phenotypic model considers only three basic types of fitness effects: the deleterious, beneficial and neutral. However, fitness effects represent a broad group forces acting on virus replication and one possible direction for further investigation would be to ungroup some of these forces and test their action. For example, on the deleterious side, the inclusion of defective interfering particles could yield another extinction mechanism, whereas on the beneficial side, the inclusion of recombination could help the viral population escape from extinction.

**Acknowledgements** LG acknowledges the support of FAPESP through the Grant Number 14/13382-1. BG and DC received financial support from CAPES.

**Author contributions** LG and DC contributed equally to this work. LMRJ and FA contributed equally to this work. Conceived the model and formulated the underlying theory: LMJR and FA. Implemented the

Table 4

| Virus | Group | $f_d$ | $d_{sp}$ | $U$ | References |
|-------|-------|-------|----------|-----|------------|
| VSV   | (-)ssRNA (V) | 0.69  | 0.50     | 0.1 | Sanjuán et al. (2004) and Furió et al. (2005) |
| Qβ    | (+)ssRNA (IV) | 0.74  | 0.52     | 0.6 | Domingo-Calap et al. (2009) and Bradwell et al. (2013) |
| TEV   | (+)ssRNA (IV) | 0.77  | 0.53     | 0.5 | Carrasco et al. (2007) and Tomas and Elena (2010) |
| Φ6    | dsRNA (III) | 0.42  | 0.34     | 0.03 | Burch and Chao (2004) and Burch et al. (2007) |
| ΦX174 | ssDNA (II) | 0.77  | 0.53     | 0.003 | Domingo-Calap et al. (2009) and Cuevas et al. (2009) |
| F1    | ssDNA (II) | 0.65  | 0.47     | 0.004 | Peris et al. (2010) and Drake (2012) |
Software Availability and Requirements The ENVELOPE program was written in C++ programming language, using the Qt 4.8.6 framework, with the Qwt 5.2.1 library. It runs on Linux and MAC-OSX operating systems and requires at least 2 GB of RAM memory and 1.5 MB of disk space. Its distribution is free to all users under the LGPL license. Binary files for Linux and MAC-OSX operating systems are available for download at: https://envelopeviral.000webhostapp.com.

Appendices

A Review of Multitype Branching Process Theory

A discrete-time multitype branching process with types or classes indexed by a nonnegative integer $r$ ranging from 0 to $R$ is described by a sequence of vector-valued random variables $Z_n = (Z_0^n, \ldots, Z_R^n)$, $(n = 0, 1, \ldots)$, where $Z_r^n$ is the number of particles of type or class $r$ in the $n$th generation. The initial population is represented by a vector of nonnegative integers $Z_0$ (also called a multi-index) which is nonzero and non-random.

The time evolution of the population is determined by a vector-valued discrete probability distribution $\xi(i) = (\xi_r(i))$, defined on the set of multi-indices $i = (i^0, \ldots, i^R)$, called the offspring distribution of the process, which is usually encoded as the coefficients of a vector-valued multivariate power series $f(z) = (f_r(z))$, called probability generating function (PGF).

The mean matrix or the matrix of first moments $M = \{M_{ij}\}$ of a multitype branching process describes how the average number of particles in each type or class evolves in time and is defined by $M_{ij} = E(Z_i^1 | Z_j^0 = 1)$, where $Z_0^i = 1$ is the abbreviation of $Z_0 = (0, \ldots, 1, \ldots, 1)$. In terms of the probability generating function $f = (f_0, \ldots, f_R)$ it is given by

$$M_{ij} = \frac{\partial f_j}{\partial z_i} (s) \bigg|_{s=1}$$  \hspace{1cm} (3)

where $1 = (1, 1, \ldots, 1)$. Typically, the mean matrix $M$ is nonnegative, and hence it has a largest nonnegative eigenvalue. When the largest eigenvalue is positive, it coincides with the spectral radius of $M$ and it is called, following Kimmel and Axelrod (2002), the malthusian parameter $\mu$.

The vector of extinction probabilities of a multitype branching process, denoted by $\gamma = (\gamma_0, \ldots, \gamma_R)$, where $0 \leq \gamma_r \leq 1$, is defined by the condition that $\gamma_r$ is the probability that the process eventually become extinct given that initially there was exactly one particle of class $r$.

The classification theorem of multitype branching processes states that there are only three possible regimes for a multitype branching process (Harris 1963; Athreya and Ney 1972; Kimmel and Axelrod 2002):

**Super-critical:** If $\mu > 1$ then $0 \leq \gamma_r < 1$ for all $r$ and, with positive probability the population survives indefinitely.

**Sub-critical:** If $\mu < 1$ then $\gamma_r = 1$ for all $r$ and with probability 1 the population becomes extinct in finite time.
Critical: If $\mu = 1$ then $\gamma_r = 1$ for all $r$ and with probability 1 the population becomes extinct; however, the expected time to the extinction is infinite.

When a multitype branching process is super-critical, it is expected that, according to the “Malthusian Law of Growth” it will grow indefinitely at a geometric rate proportional to $\mu^n$, where $\mu$ is the malthusian parameter. $Z_n \approx \mu^n W_n$ for some bounded random vector $W_n$, when $n \to \infty$. The formalization of the above heuristic reasoning is given by the Kesten–Stigum limit theorem for super-critical multitype branching processes (see Kesten and Stigum 1966a, b, 1967). If $W_n = Z_n/\mu^n$ then there exists a scalar random variable $W \neq 0$ such that, with probability one,

$$\lim_{n \to \infty} W_n = W u$$

where $u$ is the right eigenvector corresponding to the malthusian parameter $\mu$ and

$$E(W|Z_0) = v^t Z_0$$

where $v$ is the left eigenvector corresponding to the malthusian parameter $\mu$. The vectors $u$ and $v$ may be normalized so that $v^t u = 1$ and $1^t u = 1$ where $^t$ denotes the transpose of a vector. Moreover, under the assumption that $M$ is nonnegative [which is satisfied by the phenotypic model (18)], the right and left eigenvectors corresponding to the malthusian parameter are nonnegative.

The normalization of right eigenvector $u = (u_0, \ldots, u_R)$ implies that $\sum_r u_r = 1$, and therefore one has the “law of convergence of types” (see Kurtz et al. 1994)

$$\lim_{n \to \infty} \frac{Z_n}{|Z_n|} = u,$$

where $|Z_n| = \sum_r Z^n_r$ is the total population at the $n$th generation and the equality holds almost surely. Equation (6) asserts that the asymptotic proportion of a replicative class $r$ converges almost surely to the constant value $u_r$.

In particular, Eq. (6) implies that the malthusian parameter is the asymptotic relative growth rate of the population

$$\mu = \lim_{n \to \infty} \frac{|Z_n|}{|Z_{n-1}|} = \lim_{n \to \infty} \frac{1}{|Z_{n-1}|} \sum_{j=1}^{\# j} |j|$$

since $|Z_{n-1}|$ may be interpreted as the set of “parental particles” of the particles in the $n$th generation and $|Z_n|$ is the sum of the “progeny sizes” $|j|$ of the “parental particles” $j$ from the previous generation.

Now consider the quantitative random variable $\rho$ defined on the set of classes $\{0, \ldots, R\}$ and having probability distribution $(u_0, \ldots, u_R)$, called the asymptotic distribution of classes. When the classes are indexed by their expectation values, the variable $\rho$ associates to a random particle its expected class.
\[ P(\rho = r) = u_r. \]

Therefore, one can define the **average reproduction rate** of the population as

\[ \langle \rho \rangle = \sum_{r=0}^{R} r u_r. \]  

(8)

Using Eqs. (4), (5), (6) one can show that the average reproduction rate is equal to the malthusian parameter:

\[ \langle \rho \rangle = \mu. \]  

(9)

The **average population size** at the \( n \)th generation is \( |\langle Z_n \rangle| = \sum_{r=0}^{R} \langle Z_n^r \rangle \). Then for \( n \to \infty \), Eq. (4) gives \( |\langle Z_n \rangle| \approx \mu^n |\langle W_n \rangle| \approx \mu^n (W) \) and so

\[ \mu = \lim_{n \to \infty} \frac{|\langle Z_n \rangle|}{|\langle Z_{n-1} \rangle|} \]  

(10)

On the other hand, from the definition of mean matrix and its form (18), one has

\[ |\langle Z_n \rangle| = |M \langle Z_{n-1} \rangle| = \sum_{r=0}^{R} r \langle Z_{n-1}^r \rangle. \]

Now dividing by \( |\langle Z_{n-1} \rangle| \) and taking the limit \( n \to \infty \) gives

\[ \mu = \lim_{n \to \infty} \frac{|\langle Z_n \rangle|}{|\langle Z_{n-1} \rangle|} = \lim_{n \to \infty} \sum_{r=0}^{R} \frac{\langle Z_{n-1}^r \rangle}{|\langle Z_{n-1} \rangle|} = \sum_{r=0}^{R} u_r = \langle \rho \rangle \]

where here we used Eqs. (5) and (6) in the third equality from left to right.

In analogy with the characterization of the malthusian parameter as given by Eq. (7), one may define the **asymptotic populational variance**

\[ \sigma^2 = \lim_{n \to \infty} \frac{1}{|Z_{n-1}|} \sum_{j=1}^{\#[j]} \langle Z_{n-1}^j \rangle - \mu^2 \]  

(11)

and in analogy with the **mean reproduction rate**, one may define the (squared) phenotypic diversity as

\[ \sigma^2_{\rho} = \langle \rho^2 \rangle - \langle \rho \rangle^2 \]  

(12)

By decomposing the sum in Eq. (11) according to the classes \( r \), one obtains

\[ \sum_{j=1}^{\#[j]} \langle Z_{n-1}^j \rangle^2 = \sum_{r=0}^{R} \sum_{j_r=1} \#[j_r] \]
where $j_r$ runs over the particles of class $r$ for $r = 0, \ldots, R$ and $\# [ j_r ]$ are independent random variables assuming nonnegative values with probability distribution $t_r$, called fitness distribution of class $r$.

Denoting the variance of the fitness distribution $t_r$ by $\sigma_r^2$, one may write the limit in Eq. (11) as

$$
\sigma^2 = \lim_{n \to \infty} \frac{1}{|Z_{n-1}|} \sum_{j=1}^{|Z_{n-1}|} \# [ j ]^2 - \mu^2
$$

$$
= \lim_{n \to \infty} \frac{1}{|Z_{n-1}|} \sum_{r=0}^R \left[ \sum_{j=1}^{Z_{n-1}^r} \left( \frac{1}{Z_{n-1}^r} \sum_{j_r=1}^{Z_{n-1}^r} \# [ j ]^2 - r^2 \right) + Z_{n-1}^r \right] - \mu^2
$$

$$
= \lim_{n \to \infty} \frac{1}{|Z_{n-1}|} \sum_{r=0}^R (\sigma_r^2 + r^2) Z_{n-1}^r - \mu^2
$$

Then Eqs. (6), (9) and (12) give

$$
\sigma^2 = \sum_{r=0}^R (\sigma_r^2 + r^2) u_r - \mu^2 = \sum_{r=0}^R \sigma_r^2 u_r + \sigma_\rho^2
$$

(13)

The difference between the asymptotic populational variance and the (squared) phenotypic diversity, called normalized populational variance, is the weighted average of the variances of the fitness distributions

$$
\phi = \sigma^2 - \sigma_\rho^2 = \sum_{r=0}^R \sigma_r^2 u_r .
$$

(14)

In particular, when the family of fitness distributions is the deterministic family the populational variance is exactly the phenotypic diversity (that is $\phi = 0$). This is an expected result since the Delta distributions $t_r(k) = \delta_{rk}$ have zero variance, and hence the only source of fluctuation of the population size is due to its stratification into replicative classes, which is expressed by the phenotypic diversity.

B Mathematical Basis of the Phenotypic Model

Based on the general aspects of the phenomenon of viral replication described before, it is compelling to model it in terms of a branching process. At each replicative cycle, every parental particle in the replicative class $r$ produces a random number of progeny particles that is independently drawn from the corresponding fitness distribution.

A fitness distribution is a member of a location-scale family of discrete probability distributions $t_r$ parameterized by the replicative classes ($r = 0, \ldots, R$) assuming nonnegative integer values and normalized so that the expectation value of $t_r$, defined as $\sum_k k t_r(k)$, is exactly $r$ and $t_0(k) = \delta_{k0}$. Here $\delta_{kr} = 1$ if $k = r$ and $\delta_{kr} = 0$ if $k \neq r$.
$k \neq r$. Therefore, each particle in the viral population is characterized by the mean value of its fitness distribution, called *mean replicative capability*. Viral particles with replicative capability equal to zero (0) do not generate progeny; viral particles with replicative capability one (1) generate one particle on average; viral particles with replicative capability two (2) generate two particles on average, and so on. Typical examples of location-scale families of discrete probability distributions that can be used as fitness distributions are:

(a) The family of *Deterministic (Delta) distributions*: $t_r(k) = \delta_{kr}$.
(b) The family of *Poisson distributions*: $t_r(k) = e^{-r} \frac{r^k}{k!}$.

Note that in the first example, the replicative capability is completely concentrated on the mean value $r$ – that is, the particles have deterministic fitness. On the other hand, in the second example the fitness is truly stochastic.

During the replication, each progeny particle always undergoes one of the following effects:

- **Deleterious effect**: the mean replication capability of the respective progeny particle decreases by one. Note that when the particle has capability of replication equal to 0, it will not produce any progeny at all.
- **Beneficial effect**: the replication capability of the respective progeny particle increases by one. If the mean replication capability of the parental particle is already the maximum allowed, then the mean replication capability of the respective progeny particles will be the same as the replicative capability of the parental particle.
- **Neutral effect**: the mean replication capability of the respective progeny particle remains the same as the mean replication capability of the parental particle.

To define which effect will occur during a replication event, probabilities $d, b$ and $c$ are associated, respectively, with the occurrence of deleterious, beneficial and neutral effects. The only constraints these numbers should satisfy are $0 \leq d, b, c \leq 1$ and $b + c + d = 1$. In the case of in vitro experiments with homogeneous cell populations, the probabilities $c, d$ and $b$ essentially refer to the occurrence of mutations.

The probability generating function (PGF) of the phenotypic model with $b = 0$ and $t_r(k) = \delta_{kr}$ is (see Antoneli et al. (2013a, b) for details):

$$f_0(z_0, z_1, \ldots, z_R) = 1n$$
$$f_1(z_0, z_1, \ldots, z_R) = dz_0 + cz_1$$
$$f_2(z_0, z_1, \ldots, z_R) = (dz_1 + cz_2)^2$$
$$\vdots$$
$$f_R(z_0, z_1, \ldots, z_R) = (dz_{R-1} + cz_R)^R \quad (15)$$

Note that the functions $f_r(z_0, z_1, \ldots, z_R)$ are polynomials whose coefficients are exactly the probabilities of the binomial distribution $\text{binom}(k; r, 1 - d)$. The PGF in
the case with general beneficial effects and with a general family of fitness distribution
(which reduces to the previous PGF when \( b = 0 \) and \( t_r(k) = \delta_{kr} \)) is given by.

\[
\begin{align*}
f_0(z_0, z_1, \ldots, z_R) &= 1 \\
f_1(z_0, z_1, \ldots, z_R) &= \sum_{k=0}^{\infty} t_1(k) (dz_0 + cz_1 + bz_2)^k \\
f_2(z_0, z_1, \ldots, z_R) &= \sum_{k=0}^{\infty} t_2(k) (dz_1 + cz_2 + bz_3)^k \\
& \vdots \\
f_R(z_0, z_1, \ldots, z_R) &= \sum_{k=0}^{\infty} t_R(k) (dz_{R-1} + (c + b)z_R)^k \\
\end{align*}
\]

(16)

Note that in the last equation, the beneficial effect acts like the neutral effect. This is
a kind of “consistency condition” ensuring that the populational replicative capability
is, on average, upper bounded by \( R \). Even though it is possible that a parental particle
in the replicative classes \( R \) eventually has more than \( R \) progeny particles when \( t_r \)
not deterministic, the average progeny size is always \( R \).

Finally, it is easy to see that the PGF of the two-dimensional case of the phenotypic
model with \( b = 0 \) and \( z_0 = 1 \) (and ignoring \( f_0 \)) reduces to

\[
f(z) = \sum_{k=0}^{\infty} t(k) (((1 - c) + cz)^k = \sum_{k=0}^{\infty} t(k) (1 - c(1 - z))^k .
\]

(17)

This is formally identical to the PFG of the single-type model proposed by [Demetrius
et al. 1985, p. 255, eq. (49)] for the evolution of polynucleotides. In their formulation,
\( c = p^\nu \) is the probability that a given copy of a polynucleotide is exact, where
the polymer has chain length of \( \nu \) nucleotides and \( p \) is the probability of copying a single
nucleotide correctly. The replication distribution \( t(k) \) provides the number of copies
a polynucleotide yields before it is degraded by hydrolysis.

A remarkable property of the phenotypic model that was fully explored in Antoneli
et al. (2013a, b) is the fact that when \( b = 0 \) the phenotypic model is “exactly solvable”
in a very specific sense.

It is straightforward form the generating function (16), using formula (3), that the
matrix of the phenotypic model is given by

\[
M = \begin{pmatrix}
0 & d & 0 & 0 & 0 & \cdots & 0 \\
0 & c & 2d & 0 & 0 & \cdots & 0 \\
0 & b & 2c & 3d & 0 & \cdots & 0 \\
0 & 0 & 2b & 3c & 4d & \cdots & v0 \\
0 & 0 & 0 & 3b & 4c & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & Rd \\
0 & 0 & 0 & 0 & 0 & (R - 1)b & R(c + b)
\end{pmatrix}.
\]

(18)
Note that the mean matrix does depend on the fitness distributions $t_r$ only through their mean values, since $t_r$ are normalized to have the mean value $r$.

Assume for a moment that $b = 0$ (hence $c = 1 - d$). Then the mean matrix becomes upper-triangular, and hence its eigenvalues are the diagonal entries $\lambda_r = r(1 - d)$ and the malthusian parameter $\mu$ is the largest eigenvalue $\lambda_R$:

$$\mu = R(1 - d) .$$

(19)

Now suppose that $b \neq 0$ is small compared to $d$ and $c$ (hence $c = 1 - d - b$). Then spectral perturbation theory allows one to write the malthusian parameter $\mu$ as a power series

$$\mu = \mu_0 + \mu_1 b + \mu_2 b^2 + \cdots$$

where $\mu_0$ is the malthusian parameter for the case $b = 0$ and $\mu_j$ are functions of the form $R \tilde{m}_j(d, R)$. A lengthy calculation (see Antoneli et al. 2013b) gives the following result:

$$\mu = R \left( (1 - d) + (R - 1) \frac{d}{1 - d} b + O(b^2) \right) .$$

(20)

Let us return to the case $b = 0$ and consider the eigenvectors corresponding to the malthusian parameter $\mu$. The right eigenvector $\mathbf{u} = (u_0, \ldots, u_R)$ and the left eigenvector $\mathbf{v} = (v_0, \ldots, v_R)$ may be normalized so that $\mathbf{v}^\mathsf{T} \mathbf{u} = 1$ and $\mathbf{1}^\mathsf{T} \mathbf{u} = 1$, where $\mathsf{t}$ denotes the transpose of a vector. In Antoneli et al. (2013b), it is shown that the normalized right eigenvector $\mathbf{u} = (u_0, \ldots, u_R)$ is given by

$$u_r = \binom{R}{r} (1 - d)^r d^{R-r} .$$

(21)

The fact that $\mathbf{u}$ is a binomial distribution is not accidental. Indeed, it can be shown that $\mathbf{u}$ is the probability distribution of a quantitative random variable $\rho$ defined on the set of replicative classes $\{0, \ldots, R\}$, called the asymptotic distribution of classes, such that $u_r = \text{binom}(r; R, 1 - d)$ gives the limiting proportion of particles in the $r$th replicative class. Finally, when $b \neq 0$ is small, spectral perturbation theory ensures that

$$u_r = \binom{R}{r} (1 - d)^r d^{R-r} + O(b) .$$

(22)

The phenotypic model is completely specified by the choice of the two probabilities $b$ and $d$ (since $c = 1 - b - d$), the maximum replicative capability $R$ and a choice of a location-scale family of fitness distributions. Independent of the choice of family of fitness distributions, the parameter space of the model is the set $\Delta^2 \times \{ R \in \mathbb{N} : R \geq 1 \}$, where $\Delta^2 = \{(b, d) \in [0, 1]^2 : b + d \leq 1\}$ is the two-dimensional simplex (see Fig. 10).

In this parameter space, one can consider the critical curves $\mu(b, d, R) = 1$, where $\mu(b, d, R)$ is the malthusian parameter as a function of the parameters of the phenotypic model. For each fixed $R$, the corresponding critical curve is independent of the

Springer
fitness distributions and represents the parameter values \((b, d)\) such that the branching process is critical. Moreover, each curve splits the simplex into two regions representing the parameter values where the branching process is super-critical (above the curve) and sub-critical (below the curve).

One of the main results of Antoneli et al. (2013b) is a proof of the lethal mutagenesis criterion (Bull et al. 2007) for the phenotypic model, provided one assumes that all fitness effects are of a purely mutational nature. Recall that (Bull et al. 2007) assumes that all mutations are either neutral or deleterious and consider the mutation rate \(U = U_d + U_c\), where the component \(U_c\) comprises the purely neutral mutations and the component \(U_d\) comprises the mutations with a deleterious fitness effect. Furthermore, \(R_{\text{max}}\) denotes the maximum replicative capability among all particles in the viral population. The lethal mutagenesis criterion proposed by Bull et al. (2007) states that a sufficient condition for extinction is

\[
R_{\text{max}} e^{-U_d} < 1.
\]  

(23)

According to (Bull et al. 2007, 2008), \(e^{-U_d}\) is both the mean fitness level and also the fraction of offspring with no non-neutral mutations. Moreover, in the absence of beneficial mutations and epistasis (Kimura and Maruyama 1966) the only type of non-neutral mutation are the deleterious mutations. Therefore, in terms of fitness effects, the probability \(e^{-U_d}\) corresponds to \(1 - d = c\). Since the evolution of the mean matrix depends only on the expected values of the fitness distribution \(t_r\), it follows that \(R_{\text{max}}\) corresponds to \(R\). That is, the lethal mutagenesis criterion of (23) is formally equivalent to extinction criterion.
which is exactly the condition for the phenotypic model to become sub-critical. Formula (20) for the malthusian parameter provides a generalization of the extinction criterion (24) without the assumption that that all effects are either neutral or deleterious. If \( b > 0 \) is sufficiently small (up to order \( O(b^2) \)) and

\[
R \left( \frac{(1 - d) + (R - 1)}{1 - d} \right) < 1
\]

then, with probability one, the population becomes extinct in finite time.

On the other hand, a deeper exploration of the implications of nonzero beneficial effects allowed for the discovery of a non-extinction criterion. If \( b > 0 \) is sufficiently small (up to order \( O(b^2) \)), \( R \) is sufficiently large (\( R \geq 10 \) is enough) and

\[
R^3 b > 1
\]

then, asymptotically almost surely, the population cannot become extinct by increasing the deleterious probability \( d \) toward its maximum value \( 1 - b \) (see Antoneli et al. 2013b for details). In other words, a small increase in the beneficial probability may have a drastic effect on the extinction probabilities, possibly rendering the population impervious to become extinct by lethal mutagenesis (i.e., by the increase in deleterious effects).

In the theory of multitype branching processes, there are several variations as follows: continuous time, age dependent, self-regulated, etc. (see Athreya and Ney 1972; Harris 1963; Kimmel and Axelrod 2002). The implementation of a variation of the theory of multitype branching process accounting for the notions of evolutionary entropy and directionality theory (see Dietz 2005; Demetrius 2013) could be useful for studies on viral evolution. In this case, the malthusian parameter \( \mu \), which is the dominant eigenvalue of the mean matrix, could be expressed as the sum of two terms

\[
\mu = H + \Phi.
\]

The quantity \( H \) is called evolutionary entropy and \( \Phi \) is called the reproductive potential (Demetrius 2013). An interesting direction to follow would be to develop an extinction criterion based on evolutionary entropy instead of the malthusian parameter.

## C The Deterministic Selection Equation

According to (Demetrius et al. 1985; Demetrius 1985, 1987), one may associate to a multitype branching process a system of difference (or ordinary differential) equations, called selection equations, on the space of discrete probability distributions

\[
\Delta^{R+1} = \{ p \in \mathbb{R}^{R+1} : p_j \geq 0; \sum_j p_j = 1 \}
\]

over the finite state set \( \{0, \ldots, R\} \). Given a discrete multitype branching process \( Z_n \), then the expectation values \( \langle Z_n \rangle \) satisfy

\[
\langle Z_n \rangle = M^n Z_0,
\]

with \( M \) being the mean matrix of \( Z_n \). Hence \( Z_n \) is given by iteration
of the difference equation $z_n = Mz_{n-1}$. This yields a discrete-time selection equation by normalizing the difference equation, thereby obtaining

$$x_n = \frac{1}{1^tMx_{n-1}}Mx_{n-1} \quad (27)$$

where $1^t = (1, \ldots, 1)$. Then, passing (27) to continuous time one obtains a continuous-time selection equation

$$\dot{x} = [Mx - x(1^tMx)]\frac{1}{1^tMx}.$$ \hspace{1cm} (28)

Multiplying the right-hand side of Eq. (28) with the factor $1^tMx$, which is always strictly positive on $\Delta^{R+1}$, corresponds to a change in velocity (re-scaling time) and so, the solutions of (28) are the same as the solutions of

$$\dot{x} = Mx - x(1^tMx) \quad (29)$$

It follows from general considerations (see Demetrius et al. 1985; Demetrius 1985, 1987) that Eq. (29) has a unique global stable equilibrium on $\Delta^{R+1}$ given by the normalized right eigenvector $u$ of $M$ corresponding to its largest eigenvalue $\mu$. In this sense, the deterministic selection equation yields a description of the evolution of the normalized mean values of the corresponding stochastic model, thus defining a mean field (macroscopic) dynamics representing the infinite population limit of the branching process.

### D The Power Law Distribution Family

It is typical to parameterize power law distributions by the exponent $s$, which measures the “weight of the tail” of the distribution. However, we need to have a location-scale parameterized family in order to impose the same normalization as we have done for the other types of distributions. Therefore, we define the power law distribution with mean value $r$ by

$$\zeta_r(k) = \frac{(k - 1)^{s(r)}}{\zeta(s(r))}$$

for $k = 0, 1, \ldots, \infty$ and $r \geq 1$, where $\zeta(s)$ is the Riemann zeta function, defined for $s > 1$, by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

and the function $s(r)$ is given by the inverse function of

$$r = \varphi(s) = \frac{\zeta(s - 1)}{\zeta(s)} - 1.$$
Namely, \( s = \varphi^{-1}(r) \) for \( r \geq 1 \) and hence when \( 1 \leq r < \infty \) the exponent \( s \) satisfies \( 3 < s < 2 \). Moreover, the Laurent series expansion for \( r \to \infty (s \to 2) \) is given by:

\[
s(r) \approx 2 + \frac{6}{\pi^2(1 + r - C)}. \tag{30}
\]

The constant \( C \) in the previous formula is given by \( C = \left[ 6\gamma \pi^2 - 36 \zeta'(2) \right] / \pi^4 \approx 0.6974 \), where \( \gamma \) is Euler’s constant and \( \zeta'(2) \) is the derivative of \( \zeta(s) \) evaluated at 2. Observe that when the mean value \( r \geq 1 \), the exponent \( s < 3 \), and so the variance of \( z_r(k) \) is infinite.

The implementation of the pseudo-random generation of samples from the distribution \( z_r(k) \) in the ENVELOPE program is based on the algorithm of Devroye (1986) for the Zipf distribution on the positive integers, using formula 30 for the computation of the exponent \( s \) given the mean value \( r \). Pseudo-random generation for the remaining fitness distributions was implemented using the standard library of C++ programming language (this library requires C++ (2011) or superior).

### E Main Routines of the ENVELOPE Program

1:▷ Variables Defined by the User
2: Real \( b, d \); ▷ Beneficial and Deleterious Probabilities
3: Integer \( R \); ▷ Maximum Replicative Capability
4: Integer \( N \); ▷ Maximum Generation Time
5: Integer \( K \); ▷ Maximum Number of Particles
6: Integer type; ▷ Type of Fitness Distribution
7: Integer vector initial_population[0,...,\( R \)]; ▷ Initial Particle Distribution

1:▷ Global Variables
2: Integer vector malthusian[0,...,\( N \)]; ▷ Malthusian Parameter per Generation
3: Real matrix class_distribution[0,...,\( R \)][0,...,\( N \)]; ▷ Class Distribution per Generation
4: Real vector mean_rhog[0,...,\( N \)]; ▷ Average Reproduction Rate per Generation
5: Real Vector diversity[0,...,\( N \)]; ▷ Phenotypic Diversity per Generation
6: Real vector entropy[0,...,\( N \)]; ▷ Phenotypic Entropy per Generation

1:▷ Internal Variables
2: Integer \( n \); ▷ Current Generation Time
3: Integer \( T \); ▷ Current Total Progeny
4: Integer progeny; ▷ Progeny of a Replicative Class
5: Integer sampled; ▷ Random Particle Sampled
6: Integer vector particles[0,...,\( R \)]; ▷ Current Particle Distribution
7: Integer vector parents[0,...,\( R \)]; ▷ Current Parental Distribution
8: Integer vector next[0,...,\( R \)]; ▷ Next Generation Particle Distribution
9: Real effect; ▷ Random Number Between 0 and 1
1: Sample from a Fitness Distribution of type \( t \) with mean value \( m \)

2: \[\text{function } \text{FITNESSDISTRIBUTION}(\text{Real } m, \text{Integer } t)\]

3: \hspace{1em} Integer value;

4: \hspace{1em} case \( t \) do

5: \hspace{1em} \hspace{1em} \( t = 0 \): value \( \leftarrow \) (Integer) \( m \);

6: \hspace{1em} \hspace{1em} \( t = 1 \): value \( \leftarrow \) POISSON(\( m \));

7: \hspace{1em} \hspace{1em} \( t = 2 \): value \( \leftarrow \) GEOMETRIC(1/(1 + \( m \)));

8: \hspace{1em} \hspace{1em} \( t = 3 \): value \( \leftarrow \) BINOMIAL(2 \( \times \) \( m \), \( 1/2 \));

9: \hspace{1em} \hspace{1em} \( t = 4 \): value \( \leftarrow \) POWERLAW(\( m \));

10: \hspace{1em} end case

11: return value;

12: end function

1: Compute the Statistics for the Next Generation

2: \[\text{procedure } \text{STATISTICS}(\text{Integer Vector } u, \text{Integer Vector } v, \text{Integer } n)\]

3: \hspace{1em} \( \text{malthusian}[n] \leftarrow \) SUM(\( u \))/SUM(\( v \));

4: \hspace{1em} for \( i \) from 0 to \( R \) do

5: \hspace{1em} \hspace{1em} \( \text{class_distribution}[i][n] \leftarrow v[i]/\text{SUM}(v) \);

6: \hspace{1em} end for

7: \hspace{1em} \( \text{mean}_\rho[n] \leftarrow \text{AVERAGE}(\text{class_distribution}[n]) \);

8: \hspace{1em} \( \text{diversity}[n] \leftarrow \text{DIVERSITY}(\text{class_distribution}[n]) \);

9: \hspace{1em} \( \text{entropy}[n] \leftarrow \text{ENTROPY}(\text{class_distribution}[n]) \);

10: end procedure

1: Average of Class Distribution

2: \[\text{function } \text{AVERAGE}(\text{Real Vector } prob)\]

3: \hspace{1em} Real \( x = 0.0 \);

4: \hspace{1em} for \( i \) from 1 to \( R \) do

5: \hspace{1em} \hspace{1em} \( x \leftarrow x + i \times \text{prob}[i] \);

6: \hspace{1em} end for

7: \hspace{1em} return \( x \);

8: end function

1: Diversity of Class Distribution

2: \[\text{function } \text{DIVERSITY}(\text{Real Vector } prob)\]

3: \hspace{1em} Real \( x = 0.0, y = 0.0 \);

4: \hspace{1em} for \( i \) from 1 to \( R \) do

5: \hspace{1em} \hspace{1em} \( x \leftarrow x + i \times i \times \text{prob}[i] \);

6: \hspace{1em} end for

7: \hspace{1em} \( y \leftarrow \text{AVERAGE}(\text{prob}) \);

8: \hspace{1em} \( x \leftarrow x - y \times y \);

9: \hspace{1em} return \( x \);

10: end function

1: Entropy of Class Distribution

2: \[\text{function } \text{ENTROPY}(\text{Real Vector } prob)\]

3: \hspace{1em} Real \( x = 0.0 \);

4: \hspace{1em} for \( i \) from 1 to \( R \) do

5: \hspace{1em} \hspace{1em} \( x \leftarrow x + \text{prob}[i] \times \text{LOG}(\text{prob}[i]) \);

6: \hspace{1em} end for

7: \hspace{1em} return \( x \);

8: end function
1: Initialization
2: $\text{particles} \leftarrow \text{initial\_population}$;
3: $n \leftarrow 0$;
4: $\triangleright$ Main loop
5: repeat
6: $T \leftarrow \text{SUM(particles)}$;
7: $\text{parents} \leftarrow \text{particles}$;
8: $\text{next} \leftarrow [0, \ldots, 0]$;
9: for $i$ from 0 to $R$ do $\triangleright$ Generate Progeny
10: $\text{progeny} \leftarrow 0$;
11: for $j$ from 1 to $\text{parents}[i]$ do
12: $\text{progeny} \leftarrow \text{progeny} + \text{FITNESSDISTRIBUTION(Real\_i, type)}$;
13: end for
14: $\text{particles}[i] \leftarrow \text{progeny}$;
15: end for
16: if $T < K$ then $\triangleright$ Find Replicative Class of Progeny Without Cut Off
17: for $i$ from 0 to $R$ do
18: for $j$ from 1 to $\text{particles}[i]$ do
19: $\text{effect} \leftarrow \text{RANDOMREALNUMBER}(0, 1)$;
20: if $\text{effect} <= d$ then $\triangleright$ Deleterious Effect
21: $\text{next}[i-1] \leftarrow \text{next}[i-1] + 1$;
22: else if $\text{effect} > d + b$ then $\triangleright$ Beneficial Effect
23: $\text{next}[i+1] \leftarrow \text{next}[i+1] + 1$;
24: else $\triangleright$ Neutral Effect
25: $\text{next}[i] \leftarrow \text{next}[i] + 1$;
26: end if
27: end for
28: end for
29: else $\triangleright$ Find Replicative Class of Progeny With Cut Off
30: for $j$ from 1 to $K$ do
31: $\text{sampled} \leftarrow \text{RANDOMINTEGERNUMBER}(0, T)$;
32: for $i$ from 0 to $R$ do
33: if $\text{sampled} < \text{particles}[i]$ then
34: $T \leftarrow T - 1$;
35: $\text{particles}[i] \leftarrow \text{particles}[i] - 1$;
36: $\text{effect} \leftarrow \text{RANDOMREALNUMBER}(0, 1)$;
37: if $\text{effect} <= d$ then $\triangleright$ Deleterious Effect
38: $\text{next}[i-1] \leftarrow \text{next}[i-1] + 1$;
39: else if $\text{effect} > d + b$ then $\triangleright$ Beneficial Effect
40: $\text{next}[i+1] \leftarrow \text{next}[i+1] + 1$;
41: else $\triangleright$ Neutral Effect
42: $\text{next}[i] \leftarrow \text{next}[i] + 1$;
43: end if
44: break;
45: else
46: $\text{sampled} \leftarrow \text{sampled} - \text{particles}[i]$;
47: end if
48: end for
49: end for
50: end if
51: $\text{particles} \leftarrow \text{next}$; $\triangleright$ Conclude and Sumarize
52: $n \leftarrow n + 1$;
53: $\text{STATISTICS(parents, next, n)}$;
54: until $T = 0$ or $n > N$ or USERSTOP;
References

Alberch P (1991) From genes to phenotype: dynamical systems and evolvability. Genetica 84(1):5–11
Antoneli F, Bosco FAR, Castro D, Janini LMR (2013) Viral evolution and adaptation as a multivariate branching process. In: BIOMAT 2012—proceedings of the international symposium on mathematical and computational biology, vol 13. World Scientific, pp 217–243. https://doi.org/10.1142/9789814520829_0013
Antoneli F, Bosco FAR, Castro D, Janini LMR (2013) Virus replication as a phenotypic version of polynucleotide evolution. Bull Math Biol 75(4):602–628. https://doi.org/10.1007/s11538-013-9822-9
Athreya KB, Ney PE (1972) Branching processes. Springer, Berlin
Bergstrom CT, McElhany P, Real LA (1999) Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc Natl Acad Sci 96(9):5095–5100
Bradwell K, Combe M, Domingo-Calap P, Sanjuán R (2013) Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Qβ. Genetics 195(1):243–251
Bull JJ, Sanjuán R, Wilke CO (2007) Theory of lethal mutagenesis for viruses. J Virol 18(6):2930–2939. https://doi.org/10.1128/JVI.01624-06
Bull JJ, Sanjuán R, Wilke CO (2008) Lethal mutagenesis. In: Domingo E, Parrish CR, Holland JJ (eds) Origin and evolution of viruses, 2nd edn. Academic Press, London, pp 207–218. https://doi.org/10.1016/B978-0-12-374153-0.00009-6 chap. 9
Burch CL, Chao L (2004) Epistasis and its relationship to canalization in the RNA virus Φ6. Genetics 167(2):559–567
Burch CL, Guyader S, Samarov D, Shen H (2007) Experimental estimate of the abundance and effects of nearly neutral mutations in the RNA virus Φ6. Genetics 176(1):467–476
Campbell RB (2003) A logistic branching process for population genetics. J Theor Biol 225(2):195–203
Carrasco P, de la Iglesia F, Elena SF (2007) Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco Etch virus. J Virol 18(23):12979–12984
Cerf R (2015) Critical population and error threshold on the sharp peak landscape for a Moran model. Mem Am Math Soc 233(1096):1–87
Cerf R (2015) Critical population and error threshold on the sharp peak landscape for the Wright–Fisher model. Ann Appl Probab 25(4):1936–1992
Cerf R, Dalmau J (2016) The distribution of the quasispecies for a Moran model on the sharp peak landscape. Stoch Processes Appl 126(6):1681–1709
Cuesta JA (2011) Huge progeny production during transient of a quasi-species model of viral infection, reproduction and mutation. Math Comp Model 54:1676–1681. https://doi.org/10.1016/j.mcm.2010.11.055
Cuevas JM, Duffy S, Sanjuán R (2009) Point mutation rate of bacteriophage ΦX174. Genetics 183:747–749
Dalmau J (2015) The distribution of the quasispecies for the Wright–Fisher model on the sharp peak landscape. Stoch Processes Appl 125(1):272–293
Dalmau J (2016) Distribution of the quasispecies for a Galton–Watson process on the sharp peak landscape. J Appl Probab 53(02):606–613
Demetrius L (1985) The units of selection and measures of fitness. Proc R Soc Lond B 225(1239):147–159
Demetrius L (1987) An extremal principle of macromolecular evolution. Phys Scr 36(4):693
Demetrius L (2013) Boltzmann, Darwin and directional theory. Phys Rep 530(1):1–85
Demetrius L, Schuster P, Sigmund K (1985) Polynucleotide evolution and branching processes. Bull Math Biol 47(2):239–262
Devroye L (1986) Non-uniform random variate generation. Springer, Berlin
Di Mascio M, Markowitz M, Louie M, Hogan C, Hurley A, Chung C, Ho DD, Perelson AS (2003) Viral blip dynamics during highly active antiretroviral therapy. J Virol 77(22):12165–12172
Dietz K (2005) Darwinian fitness, evolutionary entropy and directional theory. BioEssays 27:1097–1101
Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. Ann Rev Microbiol 51(1):151–178
Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, Arias A (2006) Viruses as quasispecies: biological implications. In: Domingo E (ed) Quasispecies: concept and implications for viriology. Springer, Berlin, pp 51–82
Domingo E, Martínez-Salas E, Sobrino F, de la Torre JC, Portela A, Ortín J, López-Galindez C, Pérez-Breña P, Villanueva N, Núñer A (1985) The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. Gene 40(1):1–8
Domingo E, Sabo D, Taniguchi T, Weissmann G (1978) Nucleotide sequence heterogeneity of an RNA phage population. Cell 13:635–744

Domingo-Calap P, Cuevas JM, Sanjuán R (2009) The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. PLoS Genet 5(11):e1000742

Drake JW (2012) A test of Kimura’s mutation-rate conjecture. In: Mothersill CE, Korogodina VL, Seymour CB (eds) Radiobiology and environmental security. Springer, Berlin, pp 13–18

Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523

Eigen M (1993) Viral quasispecies. Sci Am 269:42–49

Eigen M, Schuster P (1979) The hypercycle. A principle of natural self-organization. Springer, Berlin

Feller W (1968) An introduction to probability theory and its applications, vol 1, 3rd edn. Wiley, New York

Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, Heldebrant C, Smith R, Conrad A, Kleinman SH, Busch MP (2003) Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 17(13):1871–1879

Fortuna MA, Zaman L, Ofria C, Wagner A (2017) The genotype-phenotype map of an evolving digital organism. PLoS Comput Biol 13(2):e1005414

Furió V, Moya A, Sanjuán R (2005) The cost of replication fidelity in an RNA virus. Proc Natl Acad Sci U S A 102(29):10233–10237

Gallant JE (2007) Making sense of blips. J Infect Dis 196(12):1729–1731

Gupta V, Dixit NM (2015) Scaling law characterizing the dynamics of the transition of HIV-1 to error catastrophe. Phys Biol 12(5):054,001

Harris TE (1963) The theory of branching processes. Springer, Berlin

Jagers P, Klebaner FC, Sagitov S (2007) On the path to extinction. Proc Natl Acad Sci U S A 104(15):6107–6111

Kesten H, Stigum BP (1966) Additional limit theorems for indecomposable multidimensional Galton–Watson processes. Ann Math Stat 37(6):1463–1481

Kesten H, Stigum BP (1966) A limit theorem for multidimensional Galton–Watson processes. Ann Math Stat 37(5):1211–1223

Kesten H, Stigum BP (1967) Limit theorems for decomposable multi-dimensional Galton–Watson processes. J Math Anal Appl 17:309–338

Kimmel M, Axelrod DE (2002) Branching processes in biology. Springer, New York

Kimura M, Maruyama T (1966) The mutational load with epistatic gene interactions in fitness. Genetics 54(6):1337

Kurtz TG, Lyons R, Pemantle R, Peres Y (1994) A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In: Athreya K, Jagers P (eds) Classical and modern branching processes, vol 84. Springer, New York, pp 181–185 IMA Vol. Math. Appl

Lambert A (2005) The branching process with logistic growth. Ann Appl Probab 15(2):1506–1535

Lee PK, Kieffer TL, Siliciano RF, Nettles RE (2006) HIV-1 viral load blips are of limited clinical significance. J Antimicrob Chemother 57(5):803–805

Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, Mullins JI (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proc Natl Acad Sci U S A 96:1492–1497

Lotka AJ (1939) Théorie analytique des associations biologiques. Part II. analyse démographique avec application particulière a’l’espece humaine. Actualités Scientifiques et Industrielles 780:123–136

Lynch M, Bürger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. J Hered 84(5):339–344

Lynch M, Gabriel W (1990) Mutation load and the survival of small populations. Evolution 44:1725–1737

Manrubia SC, Lázaro E, Pérez-Mercader J, Escarmís C, Domingo E (2003) Fitness distributions in exponentially growing asexual populations. Phys Rev Lett 90(18):188,102

Matuszewski S, Ormond L, Bank C, Jensen JD (2017) Two sides of the same coin: a population genetics perspective on lethal mutagenesis and mutational meltdown. Virus Evolut 3(1):veux004

McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF (2010) The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol 10(1):11–23

Mode CJ, Sleeman CK (2012) Stochastic processes in genetics and evolution: computer experiments in the quantification of mutation and selection. World Scientific, Singapore

Mode CJ, Sleeman CK, Raj T (2013) On the inclusion of self regulating branching processes in the working paradigm of evolutionary and population genetics. Front Genet 4:11
Nagaev AV (1967) On estimating the expected number of direct descendants of a particle in a branching process. Theory Probab Appl 12(2):314–320
Nettles RE, Kieffer TL (2006) Update on HIV-1 viral load blips. Curr Opin HIV AIDS 1(2):157–161
Nettles RE, Kieffer TL, Kwon P, Monie D, Han Y, Parsons T, Cofrancesco J, Gallant JE, Quinn TC, Jackson B (2005) Intermittent HIV-1 viremia (blips) and drug resistance in patients receiving HAART. Jama 293(7):817–829
Peris JB, Davis P, Cuevas JM, Nebot MR, Sanjuán R (2010) Distribution of fitness effects caused by single-nucleotide substitutions in bacteriophage F1. Genetics 185(2):603–609
Rong L, Perelson AS (2009) Asymmetric division of activated latently infected cells may explain the decay kinetics of the HIV-1 latent reservoir and intermittent viral blips. Math Biosci 217(1):77–87
Rong L, Perelson AS (2009) Modeling HIV persistence, the latent reservoir, and viral blips. J Theor Biol 260(2):308–331
Sanjuán R, Moya A, Elena SF (2004) The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci U S A 101:8396–8401
Schuster P, Swetina J (1988) Stationary mutant distributions and evolutionary optimization. Bull Math Biol 50(6):635–660
Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, Stern CA, Cleve JV, Yeh DJ (2014) Not just a theory: the utility of mathematical models in evolutionary biology. PLoS Biol 12(12):e1002,017. https://doi.org/10.1371/journal.pbio.1002017
Swetina J, Schuster P (1982) Self-replication with errors: a model for polynucleotide replication. Biophys Chem 16(4):329–345. https://doi.org/10.1016/0301-4622(82)87037-3
Takeuchi N, Hogeweg P (2007) Error-threshold exists in fitness landscapes with lethal mutants. BMC Evolut Biol 7(1):15
Tromas N, Elena SF (2010) The rate and spectrum of spontaneous mutations in a plant RNA virus. Genetics 185(3):983–989
Watson HW, Galton F (1874) On the probability of the extinction of families. J Anthropol Inst Great Br Irel 4:138–144
Wilke CO (2005) Quasispecies theory in the context of population genetics. BMC Evolut Biol 5(1):44
Zhu Y, Yongky A, Yin J (2009) Growth of an RNA virus in single cells reveals a broad fitness distribution. Virology 385(1):39–46. https://doi.org/10.1016/j.virol.2008.10.031

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.