A BRANCHING PROCESS FOR VIRUS SURVIVAL

J. THEODORE COX,* *Syracuse University*

RINALDO B. SCHINAZI,** *University of Colorado*

Abstract

Quasispecies theory predicts that there is a critical mutation probability above which a viral population will go extinct. Above this threshold the virus loses the ability to replicate the best-adapted genotype, leading to a population composed of low replicating mutants that is eventually doomed. We propose a new branching model that shows that this is not necessarily so. That is, a population composed of ever changing mutants may survive.

Keywords: Quasispecies; branching process; random environment; evolution

2010 Mathematics Subject Classification: Primary 60K37
Secondary 92D25

1. Introduction

Compared to other species an RNA virus has a very high mutation rate and a great deal of genomic diversity. Hence, a virus population can be thought of as an ensemble of related genotypes called quasispecies; see Eigen (1971) and Eigen and Schuster (1977). From the virus point of view, a high mutation rate is advantageous because it may create rather diverse virus genomes, which may overwhelm the immune system of the host and ensure survival of the virus population; see Vignuzzi *et al.* (2006). On the other hand, a high mutation rate may result in many nonviable individuals and hurt the quasispecies; see Sanjuan *et al.* (2004) and Elena and Moya (1999). It therefore seems that mutation rates should be high, but not too high. A simple mathematical model makes this point. Consider a virus population having genomes 1 and 2, where genome 1 has a replication rate a_1 and genome 2 has a replication rate a_2, with $a_1 > a_2$. We suppose that, when type-1 individuals replicate, the new individual has a type-1 genome with probability $1 - r$ and a type-2 genome with probability r. Type-2 genome individuals do not mutate. The model is then

$$\frac{\mathrm{d}v_1}{\mathrm{d}t} = a_1(1 - r)v_1, \qquad \frac{\mathrm{d}v_2}{\mathrm{d}t} = a_1 r v_1 + a_2 v_2, \tag{1}$$

where v_i is the number of type-i genomes for $i = 1, 2$. This is a variation of a model in Section 8.5 of Nowak and May (2000). A slightly different, but perhaps better interpretation of this model, is to think of genome 1 as being a specific (high performing) genome and genome 2 as the collection of all the other genomes in the population.

This system of differential equations is easily solved, and we can check that the ratio v_1/v_2 converges as t goes to ∞. It turns out that the limit is strictly positive if and only if $r < 1 - a_2/a_1$.

That is, in order for type 1 to be maintained in the population, the mutation r needs to be below the threshold $1 - a_2/a_1$. Hence, this model predicts that above a certain mutation threshold faithful replication of the best-adapted genotype is compromised. Moreover, there seems to be general agreement in the biology literature that above this threshold the virus population will go extinct; see Eigen (2002) and Manrubia *et al.* (2010). We propose here a simple stochastic model that shows that this is not necessarily so. In our model the population may survive, even if faithful replication of the best-adapted genotype is compromised, with the population being composed of ever changing mutants.

Our results may be biologically relevant for the following reason. An important current strategy to fight HIV and other viruses is to try to increase the mutation probability of the virus; see Eigen (2002) and Manrubia *et al.* (2010). This assumes that above a certain mutation threshold the virus will die out. Our model suggests that at least in theory this strategy may not work.

We now describe our continuous-time evolution process, which consists of individuals characterized by fitness and genotype. Let μ be a probability distribution with support contained in $[0, \infty)$, and let $r \in [0, 1]$. Start with one individual at time 0, and sample a birth rate λ from the distribution μ, which we think of as representing fitness. The individual gives birth at rate λ and dies at rate 1. Every time there is a birth, the new individual

   (i) with probability $1 - r$, keeps the same type and birth rate λ as its parent, and

  (ii) with probability r, is given a new type and a birth rate λ' sampled independently of everything else from the distribution μ.

Thinking of r as a mutation probability, we call an individual born by the latter mechanism a *mutant*. If μ has no atoms then every mutant will have a distinct birth rate. On the other hand, if μ does have atoms, a mutant may have a birth rate that has appeared previously in the population. This allows different types to have the same fitness. For convenience, we label the genotypes in the order they appear, starting with the initial individual with type 1.

Let $Z(t)$ denote the number of individuals alive at time t. We say that the evolution process *survives* if $Z(t) > 0$ for all $t \geq 0$ and *dies out* otherwise. Our main interest is in determining whether survival with positive probability is possible and by what mechanism survival can be achieved.

Theorem 1. *For $0 \leq r \leq 1$ and probability distributions μ on $[0, \infty)$, the evolution process survives with positive probability if and only at least one of the following survival conditions holds:*

$$\mu(\{\lambda : \lambda(1 - r) > 1\}) > 0, \tag{2}$$

$$\int_{\{\lambda : \lambda(1-r) \leq 1\}} \frac{\lambda r}{1 - \lambda(1 - r)} \, \mathrm{d}\mu(\lambda) > 1. \tag{3}$$

We note that the integral in (3) is infinite if μ assigns positive mass to $1/(1 - r)$. The two extreme cases, $r = 0$ and $r = 1$, are easy to understand. If $r = 0$ then (3) cannot hold and (2) reduces to $\mu([1, \infty)) > 0$. In this case, no new types are ever produced, the initial branching rate is used forever by all individuals. Conditional on the initial branching rate λ, $Z(t)$ is a linear birth–death process which survives if and only if $\lambda > 1$. Thus, (2) is equivalent to a positive probability of survival. When $r = 1$, every birth is a mutant birth, (2) cannot hold, and (3) reduces to $\int \lambda \, \mathrm{d}\mu(\lambda) > 1$. It is not hard to see that, conditional on a given individual's

branching rate λ, the total number of offspring of that individual is k with probability

$$\frac{1}{1+\lambda}\left(\frac{\lambda}{1+\lambda}\right)^{k}, \qquad k = 0, 1, \ldots,$$

and mean λ. Thus, the unconditional mean number of offspring of the first individual is $\int \lambda \, d\mu(\lambda)$, and the total number of individuals that ever live in the evolutionary process is the same as the total progeny in a Galton–Watson process with an offspring distribution which has this mean. The total progeny is infinite with positive probability if and only if this mean is larger than 1, so (3) is equivalent to a positive probability of survival.

Condition (2) corresponds to the prediction of the differential equation model (1). That is, below a certain threshold for the mutation probability, the virus can survive because a well-adapted (i.e. high λ) fixed genotype can survive. However, if (2) fails, it is still possible to have survival by (3). In this case survival holds because of a growing 'cloud' of ever changing mutants of low replicative ability.

Observe that, for any $\varepsilon > 0$ and r in [0, 1), there are distributions μ for which (2) holds but $\int \lambda \, d\mu(\lambda) < \varepsilon$. This shows that the behavior of our evolution process is drastically different from the classical Galton–Watson process in homogeneous or random environments. For these processes, survival is possible if and only if the expected offspring (or a closely related expectation) is large enough (see Harris (1989) for homogeneous environments and Smith and Wilkinson (1969) for random environments).

It is clear that if the support of μ is unbounded then (2) holds for all $r < 1$, so for interesting examples we consider distributions with compact support. Among these distributions, a natural family to consider is the uniform distribution on $[0, a]$, $a > 0$. As the following shows, this class exhibits all possible types of survival behavior depending on the exact values of a and r.

Corollary 1. *Let μ be the uniform distribution on $[0, a]$, $a > 0$. If $0 < a \leq 1$ then the evolution process dies out almost surely (a.s.) for all $r \in [0, 1]$, while if $a > 2$ then the evolution process survives with positive probability for all $r \in [0, 1]$. If $a = 2$ then the evolution process dies out a.s. for $r = 1$ and survives with positive probability for all $r \in [0, 1)$. If $1 < a < 2$ then there exists $r_c \in (1 - 1/a, 1)$ such that*

- *if $r < 1 - 1/a$ then (2) holds and the evolution process survives with positive probability,*

- *if $1 - 1/a \leq r < r_c$ then (3) holds and the evolution process survives with positive probability,*

- *if $r \geq r_c$ then the evolution process dies out a.s.*

In words, whether the population goes extinct when the mutation rate is above a certain threshold depends crucially on the value of a. If $a > 2$, there is no such threshold: the population survives for any mutation probability r. Note also that, for $1 < a < 2$, there are two distinct thresholds: $1 - 1/a$ and r_c. If $r < 1 - 1/a$, a well-adapted genome may survive forever, while if $1 - 1/a \leq r < r_c$, no fixed genome can survive forever. In this regime the population survives as a growing cloud of ever changing mutants. Finally, if $r \geq r_c$, the population goes extinct.

2. Proof of Theorem 1

Recall that we start with a single genotype-1 individual at time 0 whose λ has been sampled from distribution μ. Recall also that each birth, with probability r, produces a mutant

representing a new genotype, with its birth rate λ sampled independently from μ. Consider the branching process started by the initial type-1 individual in which all mutation births are ignored, letting X_t be the number of type-1 individuals alive at time t. Conditional on the initial branching rate λ, X_t is a birth–death process with individual birth rate $\lambda(1-r)$ and death rate 1. In particular, it is well known (see Chapter 4 of Karlin and Taylor (1975)) that it survives with positive probability if and only if $\lambda(1-r) > 1$, and that

$$\mathrm{E}(X_t \mid \lambda) = \exp((\lambda(1-r) - 1)t). \tag{4}$$

Integration of the condition $\lambda(1-r) > 1$ with respect to μ gives

$$\mathrm{P}(X_t \geq 1 \text{ for all } t > 0) > 0 \quad \text{if and only if} \quad \mu(\{\lambda : \lambda(1-r) > 1\}) > 0.$$

Now let Y_t be the number of mutants born by time t to type-1 individuals. Then $Y_t \uparrow Y_\infty$ as $t \to \infty$, the total number of mutants ever born to type-1 individuals. Note that if $r > 0$ then $Y_\infty < \infty$ if and only if $X_t = 0$ eventually. For $h > 0$, it is easy to see that

$$\mathrm{E}(Y_{t+h} - Y_t \mid \lambda, X_t) = \lambda r h X(t) + o(h) \quad \text{as } h \downarrow 0,$$

from which it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}(Y_t \mid \lambda) = \lambda r \, \mathrm{E}(X_t \mid \lambda),$$

and, therefore, using (4),

$$\mathrm{E}(Y_t \mid \lambda) = r\lambda \int_0^t \mathrm{E}(X_s \mid \lambda) \, \mathrm{d}s = r\lambda \int_0^t \exp((\lambda(1-r) - 1)s) \, \mathrm{d}s.$$

Integration with respect to the measure μ now yields

$$\mathrm{E}(Y_t) = \int_0^{+\infty} \int_0^t r\lambda \exp((\lambda(1-r) - 1)s) \, \mathrm{d}s \, \mathrm{d}\mu(\lambda).$$

By the monotone convergence theorem, $\mathrm{E}(Y_t) \uparrow \mathrm{E}(Y_\infty)$ as $t \to \infty$. Letting $m(r) = \mathrm{E}(Y_\infty)$, it is easy to show using the above that

$$m(r) = \begin{cases} +\infty & \text{if } \mu(\{\lambda : \lambda(1-r) > 1\}) > 0, \\ \displaystyle\int_{[0, 1/(1-r)]} \frac{r\lambda}{1 - \lambda(1-r)} \, \mathrm{d}\mu(\lambda) & \text{if } \mu(\{\lambda : \lambda(1-r) > 1\}) = 0. \end{cases}$$

Assume now that (2) does not hold, which implies that $X_\infty < \infty$ a.s., and, hence, $Y_\infty < \infty$ a.s. We define an auxiliary process $(Z_n)_{n \geq 0}$ similar to the tree of genotypes first introduced by Schinazi and Schweinsberg (2008) for a different model. The process (Z_n) will be a branching process with the property that $\sum_{n=1}^\infty Z_n$ is the total number of mutants that will ever appear in our evolution process. To define (Z_n), we start with the zeroth generation, the initial type-1 individual, and set $Z_0 = 1$. Next, the first generation will consist of all mutants ever born to type-1 individuals. Thus, the size of the first generation is $Z_1 = Y_\infty$. Supposing that one of these individuals has type i, we let Y_∞^i be the total number of mutants ever born to type-i individuals. Doing the same for each first generation individual generates the second generation, which has size $Z_2 = \sum_i Y_\infty^i$, where the sum is over first generation individuals.

We may continue as above, obtaining a (discrete-time) Galton–Watson process $(Z_n)_{n \geq 0}$ with offspring distribution $p_k = P(Y_\infty = k) = \int P(Y_\infty = k \mid \lambda) \, d\mu(\lambda)$, where $\sum_{n=1}^{\infty} Z_n$ is the number of mutants that will ever appear in the evolution process. The mean of the offspring distribution of $(Z_n)_{n \geq 0}$ is $m(r)$, and, hence, the total number of mutations is infinite with positive probability if and only if $m(r) > 1$.

To complete the proof, we claim that there are only two ways for the evolution process to survive: either the branching process (ignoring mutations) started by some individual survives forever with positive probability ((2) holds), or the total number of mutations is infinite with positive probability ((3) holds). It is clear that if either of these occurs then the evolution process survives with positive probability. Suppose now that both (2) and (3) fail. Then, with probability 1, each genotype that ever appears gives birth to only finitely many individuals, and also the total number of mutations is finite a.s. This means that the total number of individuals that will ever appear is finite.

3. Proof of Corollary 1

Let μ be the uniform distribution on $[0, a]$, $a > 0$. Then (2) is equivalent to $a(1 - r) > 1$. If $a(1 - r) \leq 1$ then

$$m(r) = \frac{1}{a} \int_0^a \frac{r\lambda}{1 - (1 - r)\lambda} \, d\lambda. \tag{5}$$

Case 1: $0 < a \leq 1$. Here $a(1 - r) \leq 1$ for all $r \in [0, 1]$, so (2) does not hold. Furthermore, the fact that $a \leq 1$ implies that the integrand in (5) is an increasing function of r. Thus, for all $r \in [0, 1]$,

$$m(r) \leq m(1) = \tfrac{1}{2}a < 1,$$

and, hence, (3) also fails. For every r, the evolution process dies out a.s.

Case 2: $a > 1$. A little calculus shows that

$$m(r) = -\frac{r}{1 - r} - \frac{1}{a} \frac{r}{(1 - r)^2} \ln(1 - a(1 - r)), \qquad r \in \left(1 - \frac{1}{a}, 1\right).$$

To complete the proof of Corollary 1, we will need the following properties of $m(r)$.

(P1) $m(r)$ is continuous on $(1 - 1/a, 1]$, $\lim_{r \downarrow 1 - 1/a} m(r) = \infty$, and $\lim_{r \uparrow 1} m(r) = a/2$.

(P2) If $a \geq \frac{3}{2}$ then $m(r)$ is strictly decreasing on $(1 - 1/a, 1)$.

(P3) If $1 < a < \frac{3}{2}$ then there exists $r_a \in (1 - 1/a, 1)$ such that $m(r)$ is strictly decreasing on $(1 - 1/a, r_a)$ and strictly increasing on $(r_a, 1)$.

The proof of (P1) is simple and is thus omitted. The proofs of (P2) and (P3) require some work, so we will postpone them for now and complete the proof of Corollary 1 assuming that (P2) and (P3) have been established. We consider three subcases.

- If $a \geq 2$ and $1 - 1/a < r < 1$, then, by (P2), $m(r) > m(1) = a/2 \geq 1$, so (3) holds for all $r \in (1 - 1/a, 1)$. Also, $m(1) = a/2$ implies that (3) holds for $a > 2$ but fails for $a = 2$.

- If $3/2 \leq a < 2$ then, by (P1), $m(1) < 1$, and, hence, by (P1) and (P2), there exists a unique $r_c \in (1 - 1/a, r_a)$ such that $m(r_c) = 1$. By (P2), (3) holds for $r < r_c$ but fails for $r \geq r_c$.

- If $1 < a < 3/2$ then, by (P1) and (P3), $m(r_a) < 1$. It follows that there exists a unique $r_c \in (1 - 1/a, r_a)$ such that $m(r_c) = 1$, $m(r) > 1$ on $(1 - 1/a, r_c)$, and $m(r) < 1$ on $(r_c, 1]$.

The proof of Corollary 1 is now complete except for the proofs of (P2) and (P3). At this point it is convenient to change variables. If we define the function

$$g(x) = 1 - x + \frac{1}{a}(x - x^2) \ln\left(1 - \frac{a}{x}\right), \qquad x \in (a, \infty),$$

then

$$m(r) = g\left(\frac{1}{1 - r}\right).$$

Moreover, m is increasing (decreasing) on the interval (r_1, r_2) if and only if g is increasing (decreasing) on the interval $((1 - r_1)^{-1}, (1 - r_2)^{-1})$. A little calculation gives the first three derivatives of g:

$$g'(x) = -1 - \frac{x - 1}{x - a} - \frac{1}{a}(2x - 1) \ln\left(1 - \frac{a}{x}\right),$$

$$g''(x) = -\frac{a - 3ax + 2x^2}{x(x - a)^2} - \frac{2}{a} \ln\left(1 - \frac{a}{x}\right),$$

$$g'''(x) = -\frac{a^2 + ax(2a - 3)}{x^2(x - a)^3}.$$

With some additional calculation we can explicitly check that

$$\lim_{x \downarrow a} g'(x) = -\infty, \qquad \lim_{x \to +\infty} g'(x) = 0, \tag{6}$$

$$\lim_{x \downarrow a} g''(x) = +\infty, \qquad \lim_{x \to +\infty} g''(x) = 0. \tag{7}$$

We also note that, by (P1),

$$\lim_{x \downarrow a} g(x) = \infty, \qquad \lim_{x \to +\infty} g(x) = \tfrac{1}{2}a.$$

Suppose that $a \geq \frac{3}{2}$. Then $g'''(x) < 0$ for all $x > a$, and, hence, the function g'' is strictly decreasing on (a, ∞). In view of (7), g'' must be positive on $(a, +\infty)$, which implies that g' is strictly increasing on $(a, +\infty)$. In view of (6), g' must be negative on $(a, +\infty)$, which implies that g is strictly decreasing on $(a, +\infty)$. This means that $m(r)$ is strictly decreasing on $(1 - 1/a, 1)$, so (P2) is proved.

Finally, suppose that $1 < a < \frac{3}{2}$, and set $b = a/(3 - 2a)$. Then $b > a$, $g''' < 0$ on (a, b), and $g''' > 0$ on (b, ∞). As a consequence, g'' is strictly decreasing on (a, b) and strictly increasing on (b, ∞). In view of (7) there must exist a unique $c \in (a, b)$ such that $g'' > 0$ on (a, c) and $g'' < 0$ on (c, ∞). This implies that g' is strictly increasing on (a, c) and strictly decreasing on (c, ∞). In view of (6) there must exist a unique $x_a \in (a, c)$ such that $g' < 0$ on (a, x_a) and $g' > 0$ on (x_a, ∞). This implies that g is strictly decreasing on (a, x_a) and strictly increasing on (x_a, ∞). By setting $r_a = 1 - 1/x_a$ and using the correspondence between the functions m and g, we obtain (P3).

References

EIGEN, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58,** 465–523.

EIGEN, M. (2002). Error catastrophe and antiviral strategy. *Proc. Nat. Acad. Sci. USA* **99,** 13374–13376.

EIGEN, M. AND SCHUSTER, P. (1977). The hypercycle. A principle of self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* **64,** 541–565.

ELENA, S. F. AND MOYA, A. (1999). Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. *J. Evol. Biol.* **12,** 1078–1088.

HARRIS, T. E. (1989). *The Theory of Branching Processes.* Dover Publications, New York.

KARLIN, S. AND TAYLOR, H. M. (1975). *A First Course in Stochastic Processes*, 2nd edn. Academic Press, New York.

MANRUBIA, S. C., DOMINGO, E. AND LÁZARO, E. (2010). Pathways to extinction: beyond the error threshold. *Phil. Trans. R. Soc. London B* **365,** 1943–1952.

NOWAK, M. A. AND MAY, R. M. (2000). *Virus Dynamics.* Oxford University Press.

SANJUAN, R., MOYA, A. AND ELENA, S. F. (2004). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Nat. Acad. Sci. USA* **101,** 8396–8401.

SCHINAZI, R. B. AND SCHWEINSBERG, J. (2008). Spatial and non spatial stochastic models for immune response. *Markov Process. Relat. Fields* **14,** 255–276.

SMITH, W. L. AND WILKINSON, W. E. (1969). On branching processes in random environments. *Ann. Math. Statist.* **40,** 814–827.

VIGNUZZI, M. *et al.* (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439,** 344–348