Published in final edited form as:

Evolution. 2013 June ; 67(6): 1815–1822. doi:10.1111/evo.12048.

INTER-GENERATIONAL PHENOTYPIC MIXING IN VIRAL EVOLUTION

Claude Loverdo^{1,3} and James O. Lloyd-Smith^{1,2}

¹Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

²Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

Abstract

Viral particles (virions) are made of genomic material packaged with proteins, drawn from the pool of proteins in the parent cell. It is well known that when virion concentrations are high, cells can be coinfecting with multiple viral strains that can complement each other. Viral genomes can then interact with proteins derived from different strains, in a phenomenon known as phenotypic mixing. But phenotypic mixing is actually far more common: viruses mutate very often, and each time a mutation occurs, the parent cell contains different types of viral genomes. Due to phenotypic mixing, changes in viral phenotypes can be shifted by a generation from the mutations that cause them. In the regime of evolutionary invasion and escape, when mutations are crucial for the virus to survive, this timing can have a large influence on the probability of emergence of an adapted strain. Modeling the dynamics of viral evolution in these contexts thus requires attention to the mutational mechanism and the determinants of fitness.

Keywords

Models/Simulations; Mutations; Adaptation; Extinction; Fitness

Viruses have a unique life history. Unlike most organisms, the molecules that make up viruses' "bodies" are made not only from their own genome, but also from their parent's, siblings' or neighbors' genomes. Viral particles (virions) are composed of genomic material packaged with proteins (and sometimes lipids). The proteins are drawn from a pool of proteins translated from viral genomes in the parent cell, so a genome does not necessarily match the proteins that package it. This phenotypic mixing is well recognized when a cell is coinfecting with different viruses (Závada 1976). Coinfection has even been used as a tool to characterize the effect of mutations, since viruses carrying a mutation in the same functional region cannot complement each other, and thus are of the same "complementation class" (Fields and Mahan Knipe 1990). Other classic, and extreme, examples of phenotypic mixing are the phenomena of defective viruses and interfering particles, whose genomes carry lethal genetic changes but are packaged using proteins produced by functional viruses (Aaskov et al. 2006; Pathak and Nagy 2009; Manrubia et al. 2010; Metzger et al. 2011).

Previous work has assumed that high multiplicity of infection is necessary for phenotypic mixing to occur. However, because mutation rates in viruses are high (the mean number of mutations per new genome for single stranded DNA and RNA viruses ranges at least from 0.005 to 1 (Sanjuán et al. 2010)), even a cell infected with only one virion very often contains a diversity of genomes once replication is underway. Phenotypic mixing between

³loverdo@ucla.edu.

viral generations in the absence of coinfection has been almost entirely overlooked in the literature (but see Wilke (2002); Wilke and Novella (2003)). There are many situations where multiplicity of infection is low. For many viruses, such as HIV-1 (Keele et al. 2008) or Hepatitis C (Wang et al. 2010), infections are often initiated by one or very few virions. This is particularly important in situations of evolutionary invasion, i.e. when the virus is initially not very fit in a new environment, such as a new host species or a different tissue within a host, and therefore must mutate to adapt (Antia et al. 2003; Iwasa et al. 2004; Moya et al. 2004; Pybus and Rambaut 2009). Similarly, when a virus needs to mutate to escape adaptive immunity, or a drug administered to the host, these adverse conditions may have decreased the initial viral population to levels where coinfection is negligible.

Viral evolution is an essential process in these settings. Most models of evolutionary dynamics assume a direct genotype-phenotype mapping, in the sense that when a mutation occurs, the mutant genotype is immediately associated with the mutant phenotype. We explore basic scenarios of viral replication to describe inter-generational phenotypic mixing, relate it to an established model for viral evolution, and show that it can impact evolutionary dynamics.

Model

VIRAL LIFE HISTORY

Most models of evolutionary dynamics assume perfect mapping between genotype and phenotype. If the probability for a virion to successfully infect a cell is denoted q , and the mean number of virions produced by an infected cell is denoted N (Pearson et al. 2011; Loverdo et al. 2012), this assumption means that q takes the value associated with the genome inside the virion, and N depends on the genotype of the virion which entered the cell (Fig. 1). However, the unique life history of viruses can lead to other outcomes. We first discuss when in the replication cycle mutations may occur, and then the implications for the phenotypic values N and q .

Depending on when a mutation occurs, the number of mutants produced among the offspring genomes of an infected cell and the timing of phenotypic expression of the mutation will be different. Let us discuss the case of a stamping machine replication mechanism, when the initial viral genome is used to make a template from which all new viral genomes are produced (Luria 1951). An example is retroviruses: the information from the viral RNA is incorporated into the host genomic DNA, which is then used as a template to synthesize the new viral RNA copies. A mutation that occurs during the first copying step (hereafter named an “early mutation”) makes all subsequent copies mutant, whereas a mutation occurring in the second copying step (hereafter named a “late mutation”) produces one mutant genome independently of the others. In the case of binary replication (an alternate mechanism for viral genome replication), the genome of the virion that initially infects a cell undergoes successive rounds of duplication, with each generation of new genomes being used as a template for the next (Luria 1951). In this case, a mutation in the first round of genome duplication causes half of the viral genomes produced by the cell to be mutants, approaching the case of an early mutation, whereas a mutation in the last round of duplication is equivalent to a late mutation. For simplicity, we focus on the limits of early (all genomes mutant except for the infecting genome) and late (only one mutant genome if the mutation rate is small) mutations. We will discuss the phenotypic values N^* and q^* of the “mixed” generation that occurs when an initial strain i with “pure” phenotypic parameters N_i and q_i mutates to a strain j with “pure” phenotypic parameters N_j and q_j (Fig. 1).

We first consider the number of offspring virions N^* produced by an infected cell in which a mutation from strain i to strain j occurs during viral replication. The standard genotype-phenotype mapping assumes that N^* takes the value N_j associated with the infecting genome. However, N^* depends directly on the new genomic material, as it needs to fold properly to be packaged with specific proteins to make offspring virions (D'Souza and Summers 2005; Mir et al. 2006). Proteins are also crucial, both for replication (e.g. RNA dependent RNA polymerase for RNA viruses) and as building blocks for the new virions (e.g. capsid proteins). These proteins can be produced by translation from the initial genome (e.g. for +ssRNA viruses), from the template (e.g. retroviruses or -ssRNA viruses), or from the new genomes before they are packaged (e.g. for +ssRNA viruses). For an early mutation, most of these proteins are mutant. For a late mutation, in the limit of low mutation rates, most of these proteins are of the initial type i . Thus, though the exact N^* value depends on the virus and on the mutation, N^* seems more likely to be close to N_j for an early mutation (representing a phenotypic advance relative to the standard genotype-phenotype mapping), and close to N_i for a late mutation (no shift in the genotype-phenotype mapping).

A first generation mutant virion then has a probability q^* to infect a cell and successfully initiate replication before being degraded or neutralized by the host immune response. The genome itself contributes to the success of cell infection: it may be used for translation before the first round of replication (e.g. +ssRNA virus), it needs to remain conformationally stable, and it has to interact successfully with both viral and host proteins (Wang and White 2007; Costantino et al. 2008; Gulyaev et al. 2010; Cuevas et al. 2011). Viral proteins enable the virion to interact with a susceptible cell, enter it, transport the genome to the right place within the cell, and initiate replication. An early mutation makes most proteins mutant, thus q^* likely tends to q_j , the value assumed by the standard genotype-phenotype mapping. For a late mutation, the properties of the genome are already mutant, but the proteins are mainly from the initial strain. Consequently, the value of q^* depends on the virus and the particular mutation considered. An important example is when the main effect of a mutation is to modify the viral protein that interacts with the host receptor-binding proteins. Then the phenotypic change is driven chiefly by the proteins in the coat of the new virion, and q^* tends toward the value q_i for a late mutation, i.e. the phenotypic effect of the mutation is delayed relative to the standard model.

In our analysis, we translate these biological complexities into modeling assumptions as follows. We assume that the probability that proteins are passed through several viral generations is low enough that phenotypic delay and advance act over one generation at most. When we study early mutants, we assume $q^* = q_j$ as in the standard genotype-phenotype mapping (i.e. assuming that offspring virions do not include proteins produced directly from the infecting genome) but consider N^* over a range from N_i up to N_j to represent increasing phenotypic advance. When we study late mutants, we assume $N^* = N_i$ as in the standard genotype-phenotype mapping (i.e. assuming that late mutations are sufficiently rare that they do not change the overall viral production of an infected cell) but consider q^* over a range from q_j down to q_i to represent increasing phenotypic delay. In both cases, we allow the parameters subject to phenotypic advance/delay (N^* for early mutants and q^* for late mutants) to take any value between the values for the pure phenotypes i and j . Intermediate values could represent the net effect of a mix of proteins, for example different kinds of viral polymerases determining N , or different cell receptors carried by newly produced virions determining q . For instance, if mutant proteins are translated in very low numbers after a late mutation, their effect may be not negligible if the assortment between genomes and proteins is not random, either because of how they interact with each other, or if mutant proteins are produced close in time with mutant genomes. Intermediate values could also represent a mix of phenotypic effects of the proteins and of the genome itself.

VIRAL EVOLUTIONARY DYNAMICS

If the survival of a viral lineage does not depend strongly on a set of mutations, then a minor shift in timing of the phenotypic effect of these mutations will not make a significant difference. However mutations can be crucial to the fate of a viral lineage if there are many lethal or very deleterious mutations, or if an initially unfit virus can be rescued by adaptive mutations. In the case of a deleterious mutational load high enough to cause viral extinction, if coinfection by different lineages is negligible and back or compensatory mutations are very unlikely, then the mutant lineage is doomed in any case. The exact assumptions about q and N may introduce a slight shift in the timing of extinction, but do not change the ultimate survival probability. On the contrary, for the scenario of evolutionary invasion and escape (Iwasa et al. 2004) when an unfit initial strain (basic reproductive number $R_{0,1} = q_1 N_1 < 1$) can be rescued by a mutant ($R_{0,2} = q_2 N_2 > 1$), inter-generational phenotypic mixing matters. We analyze this regime using the simple model described below.

We aim to calculate the probability of emergence, i.e. the probability that the initial viral population produces an adapted mutant leading to the survival of the viral lineage. We assume that the ratio of viruses to host target cells is low enough so that we can neglect coinfection. More broadly, we assume that host cells are present in excess and all interactions between virions are negligible. Then the probability of emergence for the founding population is the probability that at least one of the founding virions leads to emergence, computed independently for each virion. Consequently, we can focus on the case of a single founding virion without loss of generality.

We neglect back mutations from the adapted strain to the initial less fit strain, a standard assumption (Antia et al. 2003; Iwasa et al. 2004; Weissman et al. 2009). We derive rigorous results in the appendix, which coincide for small mutation rates with the more heuristic results of the main text. Despite the high overall mutation rates in viruses, very few mutations increase fitness sufficiently to allow emergence, thus the rate of appearance of such mutations, denoted μ , is very small. Additionally, we assume that an infected cell releases new virions continuously according to a budding model, so that the number of virions produced follows a geometric distribution. Under this assumption, a mutant viral lineage survives with probability $s_2 = q_2 - 1/N_2$ (Pearson et al. 2011). As discussed in the last section of the appendix, this latter assumption is unlikely to affect our results substantially.

Results

The standard approach to modeling viral evolution makes two assumptions about phenotypes after a mutation occurs from strain 1 to strain 2: the mean number of offspring virions produced by the cell where the mutation occurred is $N^* = N_1$, and the cell infection probability of the first-generation virions carrying the mutant genome is $q^* = q_2$. However, if a mutation happens early in the replication process within a cell, most genomes and most viral proteins in the infected cell are mutant, so the phenotypic effect of mutations on N^* may be advanced. Similarly, if a mutation occurs late, most genomes and proteins are of the initial type, and the phenotypic effect on q^* may be delayed. To understand how this inter-generational phenotypic mixing affects the probability of emergence of an adapted viral strain ($R_{0,2} = q_2 N_2 > 1$) from a virion of the initial unfit strain ($R_{0,1} = q_1 N_1 < 1$), we study the two cases: early mutants with $q^* = q_2$ and N^* ranging from N_1 up to N_2 for increasing phenotypic advance; and late mutants with $N^* = N_1$ and q^* ranging from q_2 to q_1 for increasing phenotypic delay. We compare the resulting probability of emergence for a given mutation rate μ .

EARLY MUTATIONS

An early mutation can occur with probability μ each time a cell is infected. The probability for the first virion to successfully infect a cell is q_1 . Then, it replicates and gives an average of $R_{0,1} = q_1 N_1$ successful offspring virions, and so on, leading to a total of

$q_1(1 + R_{0,1} + R_{0,1}^2 + \dots) = q_1 / (1 - R_{0,1})$ opportunities for mutation. When a mutation occurs, a mean of N^* offspring virions with mutant genomes is produced. If the probability for one offspring virion with mutant genome to lead to survival of the lineage is s_2 , then the probability for an early mutation to lead to survival is $N^* s_2 / (1 + N^* s_2)$ (see first section of the appendix), leading to the final probability of emergence:

$$P_{\text{emergence,early}}(N^*, q^* = q_2) = \frac{q_1 \mu N^* s_2}{(1 - q_1 N_1)(1 + N^* s_2)}. \quad (1)$$

If we compare the probability of emergence with $(N_1 \rightarrow N^* \rightarrow N_2)$ or without $(N^* = N_1)$ taking into account the phenotypic advance, we obtain the ratio:

$$\frac{P_{\text{emergence,early}}(N^*, q^* = q_2)}{P_{\text{emergence,early}}(N^* = N_1, q^* = q_2)} = \frac{N^* (1 + N_1 s_2)}{N_1 (1 + N^* s_2)}. \quad (2)$$

If the mutant strain is fitter due to better replication, $N_2 > N_1$, then phenotypic advance increases the probability of emergence compared to the standard genotype-phenotype mapping (Fig. 2a). This effect is large if $N^* \rightarrow N_2$ where $N_1 \ll N_2$ and $N_1 s_2 \ll 1$, i.e. if the mutation greatly increases an initially very low replication efficiency. An example could be a mutation conferring resistance to a drug that inhibits an enzyme involved in virion production, such as the protease inhibitor class of HIV-1 therapies (Menéndez-Arias 2010).

LATE MUTATIONS

A late mutation can occur with probability μ each time a new virion is produced. The mean number of virions produced by an initial virion is $R_{0,1} + R_{0,1}^2 + \dots = R_{0,1} / (1 - R_{0,1})$. The survival probability of a lineage initiated by a new mutant virion is $s_2 q^* / q_2$, because for the first generation after mutation, the virion may carry non-mutant proteins along with its mutant genome, so it successfully infects the next cell with probability q^* instead of q_2 . Thus the overall probability of emergence is:

$$P_{\text{emergence,late}}(N^* = N_1, q^*) = \frac{q_1 N_1 \mu s_2 q^*}{(1 - q_1 N_1) q_2} \quad (3)$$

The ratio between the emergence probabilities with $(q_1 \rightarrow q^* \rightarrow q_2)$ or without $(q^* = q_2)$ taking into account the phenotypic delay is:

$$\frac{P_{\text{emergence,late}}(N^* = N_1, q^*)}{P_{\text{emergence,late}}(N^* = N_1, q^* = q_2)} = \frac{q^*}{q_2}. \quad (4)$$

If the mutant strain is fitter due to better cell infection, $q_2 > q_1$, then this ratio is below one. If the virus is in a new tissue or a new host species and the mutation enables interaction of the virion with cell receptors (often a crucial step to change the host range (Parrish et al. 2008)), phenotypic delay may drastically reduce the probability of emergence for viruses where mutations occur late in cellular replication (figure 2b).

COMPARISON OF EARLY VS. LATE MUTATIONS

Two opposing effects contribute to the difference in emergence probabilities for early versus late mutations. Inter-generational phenotypic mixing delays phenotypic effects up to one generation for a late mutant compared to an early mutant; since evolutionary invasion and escape depend on adaptive mutations, the resulting delay in fitness gains acts to lower the probability of emergence for late mutants. At the same time, previous work has shown that early mutations yield lower emergence probabilities than late mutations because they are clustered, so their effective frequency is lower (Loverdo et al. 2012). For a given mutation rate, which gives rise to the same expected number of mutant genomes under the early and late mutation models, whether early or late mutations lead to a higher probability of emergence depends on the strength of inter-generational phenotypic mixing, which will follow from the biology of a particular virus (figure 2c). The thick line in figure 2c shows where the emergence probabilities are equal for the two types of mutation. Late mutations are more likely to emerge for scenarios close to the standard genotype-phenotype mapping, without inter-generational phenotypic mixing (bottom right corner of figure 2c). In contrast, when the phenotypic delay and advance are strongest (top left corner):

$$\frac{P_{\text{emergence}}(\text{early}, N^*=N_2, q^*=q_2)}{P_{\text{emergence}}(\text{late}, N^*=N_1, q^*=q_1)} = \frac{N_2 q_2}{N_1 q_1 (1 + N_2 s_2)} = \frac{1}{N_1 q_1} > 1, \quad (5)$$

where we used the relation $s_2 = q_2 - 1/N_2$ (Pearson et al. 2011). In this regime, the probability of emergence is larger for early mutations, reversing the conclusion of Loverdo et al. (2012). This shows that inter-generational phenotypic mixing can be a quantitatively significant effect when modeling evolutionary invasion and escape.

Discussion

Viral particles are made of a genome and proteins, produced in an infected cell. If there were several different viral genomes in the infected cell, the proteins and genome in an offspring virion may not match. While phenotypic mixing is a well recognized phenomenon at high multiplicity of infection, it is almost always overlooked in the limit of no coinfection. However, because viruses have a high mutation rate, phenotypic mixing can happen between generations in a single viral lineage. This effect is neglected in most models of viral evolution, where phenotypic effects are assumed to be expressed immediately after a mutation. To explore this phenomenon, we studied two limits: a mutation very early in the replication process, so that most genomes and thus most proteins are mutant; and a mutation at the last replication step, so that most genomes and most proteins are of the initial type. For a viral particle that successfully infects a cell with probability q and then produces an average of N new viral particles, we consider the simplified yet plausible scenarios of an early mutation advancing the phenotypic effect on N by up to one generation, and a late mutation delaying the phenotypic effect on q by up to one generation.

We focus on the probability for a beneficial mutant to emerge when a virus faces the risk of extinction because of an environmental change (adaptive immune response, drug administered to the host, new host species, new invaded tissue type, etc.). We find that the exact timing of the phenotypic effects can be crucial. In particular, a late mutation can contribute to evolutionary invasion and escape only if the first-generation mutant genome succeeds in infecting a cell despite being associated with the unfit proteins of the initial type. This scenario could describe important phenomena such as mutations affecting a viral protein that interacts with a host cell receptor (often crucial for the expansion of a virus host range (Parrish et al. 2008)) or an enzyme needed for the first steps of cell infection (such as integrases for HIV (Menéndez-Arias 2010)). Therefore this effect could slow down

emergence of new viruses or appearance of drug resistance despite the high mutation rates and large population sizes of most viruses.

We studied simplified scenarios which enabled us to discuss a broad range of phenotypic effects. First, we have studied early and late mutations separately. When replication is a stamping machine process with mutations in both steps, the contributions from mutations at each step can be added when mutations are rare, and whether phenotypic mixing increases or decreases the probability of emergence depends on the phenotypic parameters and the relative mutation rates (see appendix). Other replication processes, such as binary replication, lead to intermediate distributions of mutants and thus intermediate effects of phenotypic mixing. Another simplification in our analysis was to consider only one phenotypic parameter, q^* or N^* , which was most susceptible to being modified by phenotypic mixing for each scenario. Our results could be extended readily to any (q^*, N^*) , and the overall effect would depend on the relative values of the phenotypic parameters. Additionally, the phenotypic values for a mixed virion need not be bounded by the values of the pure types (for instance, when proteins and genome must match in order to interact successfully, in which case the mixed phenotype has lower fitness). In these circumstances, equations (1) and (3) would hold, but the direction of the change in the emergence probability would be modified. Finally, we assumed that mutations are rare. This enabled us to obtain simpler mathematical expressions. In the appendix we derive equations valid for any mutation rate if there is only one mutant strain. However at high mutation rates, there will be more than one type of mutant in an infected cell, leading to the possibility of a broader range of phenotypic values.

We focused on evolutionary invasion and escape, when the initial strain cannot survive without mutations. Inter-generational phenotypic mixing could also act on the rate of fixation of beneficial mutants in a viral population fit enough to survive. For example, a newly mutated genome encoding for an immune escape variant (i.e. a viral protein that escapes antibody recognition) could be packaged with non-mutant proteins. Consequently the phenotypic effect would be delayed, and the virion would be vulnerable to elimination by antibody-mediated host defenses, thus reducing its fixation probability. The main difference with evolutionary invasion and escape is that many cells are already infected. If cells can be superinfected, there will be additional phenotypic mixing. If not, competition for resources cannot be neglected, requiring a more detailed model of the dynamic of the susceptible cells (see first appendix of Park et al. (in press)).

Phenotypic mixing is a classic subject of study at high multiplicity of infection, but to the best of our knowledge only Wilke (2002) and Wilke and Novella (2003) have explicitly modeled the case of inter-generational phenotypic mixing in the limit of no coinfection. Wilke (2002) studied evolution of asexual organisms whose fitness depends on both their own and their parent's genome. Analyzing a quasispecies model for an infinite population, he showed that the parental fitness effect does not affect the mean fitness, but does modify the genotype frequencies. As a result, when the population is finite, the error threshold (i.e. the mutation rate at which the fittest sequence is lost) is shifted, as indicated by numerical simulations at fixed population size. Wilke and Novella (2003) presented a deterministic analysis of the steady state of a system corresponding to a late mutation with the effects on q delayed by one generation. They showed that in the limit of a very deleterious mutant, phenotypic mixing doubles the proportion of mutant genomes, because mutant genomes can replicate before the lethal phenotype is expressed. They focused on the proportion of mutant genomes at the mutation-selection balance for a very large viral population size. Our study complements these two earlier studies, by examining the stochastic dynamics of a population of varying size, where the outcome of interest is whether the viral population survives or not.

In our model, one mutation is enough to bring a substantial increase in fitness, whereas sometimes several mutations are needed (Shih et al. 2007). If the intermediate mutants have intermediate fitness, the reasoning for each step stems directly from our results for a single step. At first glance, it seems that phenotypic delay might assist with crossing fitness valleys (i.e. when the intermediate mutant has low fitness) (Bonhoeffer et al. 2004; Weinreich and Chao 2005; Weissman et al. 2009), since the deleterious effect of the intermediate genome can be masked by phenotypic mixing with the initial strain. However, this same principle means that the beneficial effects of the final mutant (the “far side of the valley”) will be masked by the deleterious intermediate, so the only difference will be a slight shift in the timing of when the deleterious phenotype is expressed. Thus, in general, inter-generational phenotypic mixing does not help in crossing fitness valleys.

Our results also have implications for basic virology research. Phenotypes are often used to measure mutation rates. A common approach is to let a virus replicate for a known number of generations under conditions where mutations at a specific target are neutral, then viruses are transferred to new culture conditions where only specific mutants can survive, and plaques are counted to infer the mutation rate (la Torre et al. 1990; Chao et al. 2002; Malpica et al. 2002; Raney et al. 2004). This implicitly assumes the standard genotype-phenotype mapping, and thus inter-generational phenotypic mixing may bias the resulting mutation rate estimate. Holland et al. (1989) noticed that such estimates of mutation rates vary widely with multiplicity of infection and proposed a solution. In their system, antibodies inactivate viral particles displaying a wild-type surface protein. They apply antibodies to the plate only after a time sufficient for cell infection in the new medium, so as to delay selection enough that it acts on beyond the first mutant generation, which no longer exhibit phenotypic mixing. Our results show that similar precautions may be needed to measure the mutation rate accurately even at very low multiplicity of infection. Conversely, this framework could be used to study inter-generational phenotypic mixing directly. If the experimental conditions are chosen to ensure negligible levels of coinfection, then the ratio of the number of plaques under immediate versus delayed selection will quantify the effect of inter-generational phenotypic mixing for the system in question. This antibody-binding approach would characterize the effect on q^* , but related methods could be devised to assess the impact on N^* . Applying this approach to a range of host-virus systems would show the extent of deviations from the standard genotype-phenotype mapping, which in conjunction with our results will indicate the potential importance of this phenomenon for viral evolution.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank S. J. Schreiber and M. Park for providing useful comments on earlier drafts of this manuscript, and A. Siddiqui for providing valuable context over a breakfast table. Both authors are supported by the NSF grant EF-0928690. JLS is grateful for the support of the De Logi Chair in Biological Sciences, and the RAPIDD program of the Science & Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health.

LITERATURE CITED

- Aaskov J, Buzacott K, Thu HM, Lowry K, Holmes EC. Long-term transmission of defective RNA viruses in humans and *Aedes* mosquitoes. *Science*. 2006; 311:236–238. [PubMed: 16410525]
- Antia R, Regoes RR, Koella JC, Bergstrom CT. The role of evolution in the emergence of infectious diseases. *Nature*. 2003; 426:658–661. [PubMed: 14668863]

- Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. Evidence for positive epistasis in HIV-1. *Science*. 2004; 306:1547–1550. [PubMed: 15567861]
- Chao L, Rang CUC, Wong LEL. Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage ϕ 6. *J Virol*. 2002; 76:3276. [PubMed: 11884552]
- Costantino, Da; Pflingsten, JS.; Rambo, RP.; Kieft, JS. tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Biol*. 2008; 15:57–64. [PubMed: 18157151]
- Cuevas JM, Domingo-Calap P, Sanjuán R. The Fitness Effects of Synonymous Mutations in DNA and RNA Viruses. *Mol Biol Evol*. 2011; 29:17–20. [PubMed: 21771719]
- D'Souza V, Summers MF. How retroviruses select their genomes. *Nat Rev Microbiol*. 2005; 3:643–655. [PubMed: 16064056]
- Fields, BN.; Mahan Knipe, D. *Fields virology*. Raven Press; 1990.
- Gultyaev AP, Fouchier RaM, Olsthoorn RCL. Influenza virus RNA structure: unique and common features. *Int Rev Immunol*. 2010; 29:533–556. [PubMed: 20923332]
- Harris, TE. *The theory of branching processes*. Dover Phoenix Editions; Mineola, NY: 1963.
- Holland J, Torre JDL, Steinhauer D, Clarke D, Duarte E, Domingo E. Virus mutation frequencies can be greatly underestimated by monoclonal antibody neutralization of virions. *J Virol*. 1989; 63:5030. [PubMed: 2479770]
- Iwasa Y, Michor F, Nowak MA. Evolutionary dynamics of invasion and escape. *J Theor Biol*. 2004; 226:205–214. [PubMed: 14643190]
- Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson Ja, Ping L-H, Swanstrom R, Tomaras GD, Blattner Wa, Goepfert Pa, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci USA*. 2008; 105:7552–7557. [PubMed: 18490657]
- Loverdo C, Park M, Schreiber SJ, Lloyd-Smith JO. Influence of viral replication mechanisms on within-host evolutionary dynamics. *Evolution*. 2012; 66:3462–3471. [PubMed: 23106710]
- Luria, S. *Cold Spring Harbor symposia on quantitative biology*. Vol. 16. Cold Spring Harbor Laboratory Press; 1951. The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction; p. 463
- Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, García-Arenal F. The rate and character of spontaneous mutation in an RNA virus. *Genetics*. 2002; 162:1505–1511. [PubMed: 12524327]
- Manrubia SC, Domingo E, Lázaro E. Pathways to extinction: beyond the error threshold. *Philos Trans R Soc Lond B*. 2010; 365:1943–1952. [PubMed: 20478889]
- Menéndez-Arias L. Molecular basis of human immunodeficiency virus drug resistance: an update. *Antiviral Res*. 2010; 85:210–231. [PubMed: 19616029]
- Metzger VT, Lloyd-Smith JO, Weinberger LS. Autonomous targeting of infectious superspreaders using engineered transmissible therapies. *PLoS Comput Biol*. 2011; 7:e1002015. [PubMed: 21483468]
- Mir MA, Brown B, Hjelle B, Duran WA, Panganiban AT. Hantavirus N protein exhibits genus-specific recognition of the viral RNA panhandle. *J Virol*. 2006; 80:11283–11292. [PubMed: 16971445]
- Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. *Nat Rev Microbiol*. 2004; 2:279–288. [PubMed: 15031727]
- Park M, Loverdo C, Schreiber SJ, Lloyd-Smith JO. Multiple scales of selection influence the evolutionary emergence of novel pathogens. *Phil Trans R Soc B*. in press.
- Parrish CR, Holmes EC, Morens DM, Park E-C, Burke DS, Calisher CH, Laughlin Ca, Saif LJ, Daszak P. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol Mol Biol R*. 2008; 72:457–470.
- Pathak KB, Nagy PD. Defective Interfering RNAs: Foes of Viruses and Friends of Virologists. *Viruses*. 2009; 1:895–919. [PubMed: 21994575]

- Pearson JE, Krapivsky P, Perelson AS. Stochastic Theory of Early Viral Infection: Continuous versus Burst Production of Virions. *PLoS Comput Biol*. 2011; 7:e1001058. [PubMed: 21304934]
- Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Microbiol*. 2009; 10:540–550.
- Raney JL, Delongchamp RR, Valentine CR. Spontaneous mutant frequency and mutation spectrum for gene A of phiX174 grown in *E. coli*. *Environ Mol Mutagen*. 2004; 44:119–127. [PubMed: 15278916]
- Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. *J Virol*. 2010; 84:9733–9748. [PubMed: 20660197]
- Shih ACC, Hsiao TC, Ho MS, Li WH. Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc Natl Acad Sci USA*. 2007; 104:6283–6288. [PubMed: 17395716]
- la Torre JD, Wimmer E, Holland JJ. Very high frequency of reversion to guanidine resistance in clonal pools of guanidine-dependent type 1 poliovirus. *J Virol*. 1990; 64:664. [PubMed: 2153232]
- Wang GP, Sherrill-Mix Sa, Chang K-M, Quince C, Bushman FD. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J Virol*. 2010; 84:6218–6228. [PubMed: 20375170]
- Wang S, White KA. Riboswitching on RNA virus replication. *Proc Natl Acad Sci USA*. 2007; 104:10406–10411. [PubMed: 17553958]
- Weinreich DM, Chao L. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution*. 2005; 59:1175–1182. [PubMed: 16050095]
- Weissman DB, Desai MM, Fisher DS, Feldman MW. The rate at which asexual populations cross fitness valleys. *Theor Popul Biol*. 2009; 75:286–300. [PubMed: 19285994]
- Wilke CO. Maternal effects in molecular evolution. *Phys Rev Lett*. 2002; 88:18–21.
- Wilke CO I, Novella S. Phenotypic mixing and hiding may contribute to memory in viral quasispecies. *BMC Microbiol*. 2003; 3:11. [PubMed: 12795816]
- Závada J. Viral pseudotypes and phenotypic mixing. *Arch Virol*. 1976; 50:1–15. [PubMed: 816332]

Figure 1.

Viral replication: when a mutation occurs, inter-generational phenotypic mixing blurs the genotype-phenotype mapping. The details may vary but this cartoon gives a broad idea of the limiting cases. The standard genotype-phenotype mapping implies that the probability for a virion to infect a cell (q) takes the value associated with the genome inside the virion, and that the mean number of virions produced by an infected cell (N) depends on the genotype of the virion which entered the cell. An early mutation will make most proteins mutant, while a late mutation will produce few if any mutant proteins. Consequently, the number of virions produced by a cell where a mutation in viral replication occurs (N^*) and the probability for newly produced virions carrying a mutant genome to successfully infect the next cell (q^*) may differ from the standard assumption.

Figure 2.

How the probability of emergence is changed by inter-generational phenotypic mixing, (a) Probability of emergence when mutations occur early in viral replication. In this plot $q_{\text{early}}^* = q_2$ (as assumed in the standard genotype-phenotype mapping), and N_{early}^* ranges from N_1 up to N_2 to represent phenotypic advance. The three lines show different parameters describing the cell infection probability q for each strain. (b) Probability of emergence when mutations occur late in viral replication. In this plot $N_{\text{late}}^* = N_1$ (as assumed in the standard genotype-phenotype mapping), and q_{late}^* ranges from q_2 down to q_1 to represent phenotypic delay. The three lines show different scenarios for the mean number of virions produced, N , for each strain. (c) Ratio of the probability of emergence in the case of early and late mutations $P_{\text{emergence,early}}(N_{\text{early}}^*, q_{\text{early}}^* = q_2) / P_{\text{emergence,late}}(N_{\text{late}}^* = N_1, q_{\text{late}}^*)$ as a function of q_{late}^* for the late mutant and N_{early}^* for the early mutant. This plot shows the combined effects of inter-generational phenotypic mixing (which decreases the probability of emergence for late mutations, relative to early mutations, because the beneficial phenotypes are delayed) and correlations among mutations (which increases the probability of emergence for late mutations, relative to early mutations). For all plots, $\mu = 0.001$.