Vol. 82, No. 9

Rapid Adaptive Amplification of Preexisting Variation in an RNA Virus[▽]

Ranendra N. Dutta,[1]† Igor M. Rouzine,[2]† Sarah D. Smith,[1] Claus O. Wilke,[3] and Isabel S. Novella[1]*

Department of Medical Microbiology and Immunology, College of Medicine, University of Toledo, Toledo, Ohio 43614[1]; Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts 02111[2]; and Center for Computational Biology and Bioinformatics, Section of Integrative Biology, and Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, Texas 78712[3]

The amount and nature of preexisting variation in a population of RNA viruses is an important determinant of the virus's ability to adapt rapidly to a changed environment. However, direct quantification of this preexisting variation may be cumbersome, because potentially beneficial alleles are typically rare, and isolation of a large number of subclones is required. Here, we propose a simpler method. We infer the initial population structure of vesicular stomatitis virus (VSV) by fitting a mathematical model of asexual evolution to an extensive set of measurements of VSV fitness dynamics under various conditions, including new and previously published data. The inferred variation of fitness in the initial population agrees very well with the results of direct experiments with subclone fitness quantification. From the same procedure, we also estimate the mean fitness effect of beneficial mutations (selection coefficient s), the percentage of sites in the genome that are under moderate positive or negative selection, and the percentage of sites where beneficial mutations may potentially occur. For VSV strain MARM U evolving in BHK-21 cells, the three parameters have values of 0.39, 9%, and 0.06%, respectively. The method can be generalized and applied easily to other rapidly evolving microbes, including both asexual microorganisms and those with recombination.

RNA viruses have an error-prone replication machinery and form highly polymorphic populations (7–9). As a result, RNA viruses exposed to a new or altered environment frequently respond with rapid adaptation to the new conditions. This high evolutionary potential is one of the major obstacles to the development of effective antiviral treatments, and it makes RNA viruses prime candidates for newly emerging or reemerging diseases.

The quick response of RNA viruses to new environmental conditions is explained by the combination of large population sizes, high mutation rates (10^{-4} to 10^{-5} per base per round of copying), rapid population turnover, and in the case of segmented and positive-stranded viruses, effective mechanisms of recombination. The adaptive response is especially swift when it is caused by the selective amplification of preexisting variants rather than by the de novo generation of beneficial mutations. For example, resistance to monotherapy against human immunodeficiency virus type 1 (HIV-1) infection can arise in as little as 1 to 2 weeks (30, 36), because mutants that confer resistance exist, at a small frequency, prior to therapy (4, 17, 20). Thus, while long-term adaptation is caused by a gradual accumulation of new beneficial alleles due to mutation and, for some viruses, yeast, and bacteria, recombination (4, 5, 14, 18, 28, 31, 33, 35, 38), the short-term adaptive response is determined by the preexisting variation that is present in a viral population of a given size and by the proportion of these mutants that are beneficial under specific environmental conditions. Consequently, in order to understand (and possibly be able to pre-

dict) viral adaptation to new environmental conditions, we need to know the frequency and nature of preexisting mutants in a viral population.

In principle, the question about standing variation can be answered directly. One only has to isolate a large number of clones from a viral population, sequence them, and determine their fitness in the environment of interest. However, in practice this approach may be not feasible, because even very rare beneficial mutants are potentially important. When strong selective conditions (e.g., drug therapy) are imposed, these rare sequences will rapidly rise in frequency and take over the population. Therefore, to assess the frequency of some very rare beneficial mutants in the initial population, we would have to characterize at least hundreds of thousands of isolates. Here, we take a different approach. We infer the initial state of a viral population by fitting a mathematical model of virus growth and evolution to a large set of independent measurements of the population dynamics of vesicular stomatitis virus (VSV). Our main data set consists of measurements of the initial speed of adaptation of a VSV laboratory strain during replication in BHK-21 cells. We expanded this data set to include the published measurements of the VSV equilibrium fitness after prolonged passaging (25), the speed at which the equilibrium is reached (22), the relationship between viral fitness and the critical bottleneck size at which continued passaging does not lead to fitness changes (24), and the clonal analysis of a mutant population (11). We show that we can fit a simple model of asexual replication with a single set of six parameters fitting to the entire set of measurements. To our knowledge, a fit encompassing such a wide array of different aspects of viral population dynamics has not previously been carried out. From our fitting, we obtain not only the amount of preexisting variation in the initial population but also the estimates for the mean effect of a beneficial mutation of fitness,

---

* Corresponding author. Mailing address: Department of Medical Microbiology and Immunology, University of Toledo Health Science Campus, Toledo, OH 43614. Phone: (419) 383-6442. Fax: (419) 383-4221. E-mail: isabel.novella@utoledo.edu.
† These authors contributed equally to the work.
▽ Published ahead of print on 20 February 2008.

the number of sites where such mutations can potentially occur, and the total number of sites in the genome that are under moderate selection, relevant over the time scale of 10 to 100 passages.

Our computational model naturally accounts for the effects of linkage (inheritance of genomic sites together, as a set) that are known to slow evolution down significantly and cause Muller's ratchet in small populations. The basic one-locus model (reviewed in reference 32) generally does not apply to the evolution of finite asexual populations. Instead, we can describe multisite evolution accurately by using a recently developed analytic approach that is based on the concept of moving fitness distributions (18, 31, 33, 35; but see Gerrish et al. [13]). We found that the analytic approach, although valid on any time scale, is more convenient in the case of long-term evolution, when the fitness distribution profile and the adaptation rates are independent of the initial conditions. For the (mostly) short-range evolution studied here, we employed instead direct Monte Carlo simulation of the model.

MATERIALS AND METHODS

Cells and viruses. We propagated BHK-21 hamster kidney cells from John Holland's laboratory in Eagle's minimal essential medium (MEM) containing 7% heat-inactivated (60°C, 30 min) bovine calf serum. For virus passages and competitions, we used fetal bovine serum. For the growth of BHK-21 cells, we supplemented the medium with 0.06% proteose peptone no. 3 (Difco). We used mouse hybridoma cells, kindly provided by Douglas Lyles (19), for the production of I1 antibody as described previously (10).

Wild-type (wt) VSV (Mudd-Summers strain; Indiana serotype) was the progenitor of all other strains, and we used it as the reference for fitness determinations. The other viral strains employed here were monoclonal antibody (MAb)-resistant mutants (MARMs) with a single amino acid substitution, Asp→Ala, at amino acid 259 (or Asp→Gly at amino acid 257) of the G glycoprotein that confers resistance to I1 MAb. MARM U is a clone isolated from the wt in the presence of MAb I1 and amplified once at a low multiplicity of infection (MOI) in BHK-21 cells. Its sequence differs only from that of the wt in the I1 resistance marker, and its relative fitness is 1.0 ± 0.2 (24). MARM X had a relative fitness of 3.05 ± 0.03, and it was the progeny of MARM U after 20 plaque-to-plaque passages, followed by 61 large population passages in BHK-21 cells (24). For this study, we generated strain MARM U35 by carrying out 35 large population passages of MARM U in BHK-21 cells, and its fitness is 5.5 ± 0.3. MARM C had a fitness of 0.93 ± 0.02 and it was the progeny of MARM U after 20 plaque-to-plaque passages, followed by two large population passages (22). MARM N had a fitness of 0.38 ± 0.01 (24), and it differs from the wt in the resistance marker and a nucleotide substitution in the polymerase gene (nucleotide 8700) (26). For the present studies, we carried out an additional amplification passage of the wt, MARM U, and MARM N that resulted in no changes in the relative fitness of MARM U (1.00 ± 0.05) (21) and a small increase in the relative fitness of MARM N to 0.43 ± 0.02 (26). Thus, compared to previously published data, for the present work, we used strains with one additional amplification passage for the fitness determinations of MARM U and derived populations, MARM N and derived populations, and MARM U35 and derived populations.

Virus passages. We used flasks with BHK-21 cell monolayers at a density of about 10^5 cells/cm^2 to infect with MARM U at various transmission population size (N_t) values, ranging from 4 to 7.5×10^5 PFU per passage. We used different flask sizes to maintain the plaque density and a constant MOI. After cells were inoculated, we incubated the cell monolayers for 10 min at room temperature for attachment and for 45 additional minutes at 37°C for penetration, and then we added MEM (200 μl/cm^2). We then incubated the flasks at 37°C for 20 to 22 h, following which, we harvested the virus, diluted it, and used it to infect fresh BHK-21 monolayers. This process was continued for 20 consecutive passages, and we determined the fitness at every fourth passage. In a first set of experiments, we analyzed N_t values between 4 and 3,000, and we added agarose to the MEM overlay; we recovered progeny from plaques (about 10^7 PFU/plaque). These infections were carried out with an initial density of 10 PFU/cm^2 or less (MOI, <0.0001 PFU/cell). For an N_t of 50 PFU or less, we randomly picked and pooled individual plaques from T25 flasks. For an N_t between 100 and 3,000

PFU, we infected cell monolayers at different dilutions, recovered the full content of the overlay in each flask, and counted the plaques at each dilution. To maintain the MOI at an approximately constant rate for this range of population sizes, we increased the number of cells and used larger T75 flasks for an N_t of 500, and we sampled and combined the content of several flasks for the larger N_t value (two flasks for an N_t of 1,000 and four to six flasks for an N_t of 3,000). We then used the recovered MEM overlay that contained the desired number of plaques (within a 10% error) for subsequent passages. We carried out five replicas for N_t values of 4, 8, and 15, and three replicas for all other N_t values. In a second set of experiments, we performed passages with N_t values between 480 and 7.5×10^5 at a constant MOI of 0.015 PFU/cell, using flask sizes that ranged between 96-well plates (0.32 cm^2/well) and 500-cm^2 plates. In this case, we did not add agarose to the MEM, and we allowed the infections to proceed until the cytopathic effect was complete, and the maximum titer was reached (10^{10} PFU/ml). For the next passage, we diluted the progeny as needed to maintain a constant N_t. In both sets of experiments, infections start with individual virions targeting individual cells that represent a small fraction of the total number of cells available in the monolayer. At 4 h postinfection, progeny release starts and continues for the next 5 to 6 h, until each cell has generated over 10^4 PFU. In the absence of agarose (e.g., in liquid medium), the new progeny will diffuse freely and infect the remaining cells. We refer to this regimen as the "dilution" passages, and viral progenies at the end of a passage are the result of two rounds of infection. In contrast, in the presence of agarose, the diffusion of virions is limited, and only cells in close proximity to the initially infected cell are available for infection. We refer to this regimen as "plaque pooling," and progenies are the product of three to four rounds of infection. Note that up to three rounds of infection per passage occur at high MOI in these regimens. However, three rounds of infection at high MOI are insufficient for the accumulation of a significant amount of defective interfering particles that could otherwise confound our results. Even during undiluted passages, when the MOI can be as high as 1,000 PFU/cell, we did not observe any sign of interference before passage five (I. S. Novella, unpublished results).

We define the critical bottleneck size as the N_t value at which the overall fitness of a strain is the same for the initial population and after 20 passages. We determined the critical bottleneck size for MARM U35 by carrying out six replicas of 20 passages at an N_t of 300 with a regimen of plaque pooling. We then determined the fitness of the evolved populations. In addition, we carried out six replicas of passages at an N_t of 30, which was predicted to result in Muller's ratchet operation and overall fitness loss.

We carried out six replicas of two plaques-to-two plaques passages of MARM N for 20 passages. We determined the fitness of all the replicas at the end of the experiment, as well as the fitness at passages 1, 2, 3, 5, 10, and 15 for two of the replicas.

Fitness determination and calculation of speed of adaptation. We measured the fitness of evolved MARM populations in competition against the wt reference strain, as described previously (16). For the fitness determinations of MARM U progeny, we mixed test virus with the wt, and used the mixture to infect BHK-21 cell monolayers at an MOI of 0.1. The initial ratio of wt versus test virus was determined by performing triplicate plaque assays with and without I1. After 20 to 24 h of incubation, we diluted the viral yield to start a second competition passage, and we determined the new wt/MARM virus ratios by plaque assay. We carried out up to four competition passages per fitness determination. We then log transformed the wt/MARM virus ratios, plotted them against the passage number, did a linear regression of the log-transformed ratios against the passage number, and determined the slope, m, of the regression line. Finally, we calculated the fitness, w, as given by the exponential expression $w = \exp(m)$. For fitness determinations of the progeny of MARM N and MARM U35, we used three single-passage competitions and calculated fitness as the average of the three changes in ratio.

We measured the speed of adaptation, V, for a given replica by calculating the natural logarithm for all fitness values in the replica and then fitting the function Vt (where t is time measured as the passage number) to the log-transformed fitness values.

Computational model. We used a simple model of an asexual population of the Wright-Fisher type with discrete nonoverlapping virus generations (33). A genome consists of L sites, each of which can be in either of two states (alleles), beneficial or deleterious. A beneficial allele makes a fitness contribution of 1 to the fitness of the genome, whereas a deleterious allele makes a fitness contribution of $\exp(-s)$. The total fitness of the genome is the product of all the fitness contributions of the individual sites, i.e., epistasis is neglected. Hence, a genome with k deleterious alleles leaves progeny whose average number is $\exp[-(k - k_{av})s]$, where $\exp(-k_{av}s)$ is the average fitness of the population (i.e., the k_{av} value is somewhat smaller than the average value of k). We measured fitness in

relative units, with respect to our reference wt strain, whose fitness is set to 1. To relate absolute fitness values determined with respect to the maximum possible fitness in the cell culture $\exp(-k_{av}s)$ to relative fitness values, we introduced an additional parameter, the number of sites carrying deleterious alleles in the reference strain, k_{ref}. These are the sites at which beneficial mutations can potentially occur.

The fitness diversity of the initial population of a strain is characterized, in our analysis, by the variance of the average number of deleterious alleles between individual genomes, which we denote by $d = \text{Var}[k]_{t=0}$. The initial distribution of the deleterious allele number, k, among genomes is assumed to be Gaussian. The model is asexual and does not include recombination, which has not been observed for VSV (29). The actual progeny number of a class of genomes with the same fitness fluctuates according to the Poisson distribution, subject to the restriction that the total population remains constant and equal to N. An allele at a site can mutate to the opposite allele with the probability μ per generation. The total genomic mutation rate and the total genomic rate of beneficial mutations are denoted by $U = \mu L$ and $U_b = \mu k$, respectively.

We equate the effective population size, N, used in the simulations with the size of the infecting population, N_t. For a population with a time-dependent size that expands exponentially during each passage, the effective population size is calculated as a harmonic mean, which is dominated by the smallest population size at each passage. Therefore, the population size is associated with the first round of infection in a passage.

A second parameter of the model, the selection coefficient s, is also associated with the first round of infection in a passage. The amount of progeny from the first round of infection is enough to result in widespread coinfection in subsequent rounds in the same passage. A high degree of coinfection leads to complementation of gene products of different genetic variants within the same cell, effectively suppressing selection. Thus, under widespread coinfection, the phenotype fitness of virions is similar to the population average, in contrast to the variable fitness of their RNA genomes (26, 39). Because both N and s are associated with the first round of infection in a passage, the effective number of virus generations in our simulations is equal to the number of passages in experiments.

By contrast, mutations occur in every round of infection, and the number of rounds of infection depends on the details of the passage regimen (see "Virus passages" above). Therefore, we increase the mutation rate per round of infection by a factor of three for plaque-pooling passages and by a factor of two for dilution passages, and denote these mutation rates by $\mu_P = 3\mu$ and $\mu_D = 2\mu$, respectively.

To simulate the stochastic evolution of a population of constant size, we used a Monte Carlo algorithm written with MATLAB software, similar to that used previously (31, 33). We classified all genomes according to the number of deleterious alleles, k (i.e., according to fitness), and monitored the population size of each fitness class in consecutive discrete generations. All fitness classes containing more than 100 genomes were treated deterministically, neglecting random genetic drift. The other nonempty classes were considered stochastic. We checked that increasing the boundary between deterministic and stochastic treatment above 100 did not change results noticeably. The progeny number for a deterministic fitness class was replaced with the average progeny number of that class. For stochastic classes, if their combined size did not exceed 30% of an entire population, we used a pseudorandom number generator to simulate a Poisson distribution around the average progeny number of each class. Then, the sizes of all classes were renormalized to ensure that the total population size did not change and rounded to the nearest integer value. If the total fraction of stochastic classes was larger than 30%, because the total population was small, all fitness classes in a population were considered stochastic. In this case, we used a broken-stick algorithm to simulate a polynomial distribution of progeny among classes, as follows. We represented each class with an interval of the length proportional to the product of its current size and fitness and randomly distributed N values between the intervals. The number of points that hit an interval was the progeny number in the class.

Because the relevant values of both genomic rates U and U_b are considerably less than 1 (see below), we allowed only one mutation per genome for both deleterious and beneficial mutations. We calculated the number of mutant genomes with $k - 1$ or $k + 1$ alleles originating from a class with k alleles, for a deterministic class, as the corresponding average value, or for a stochastic class, as a random Poisson value with the corresponding average. The resulting numbers of mutant genomes were transferred between adjacent fitness classes. We monitored the average fitness of a population as a function of time.

Fitting model parameters to the data. The average mutation rate per site (nucleotide) per allele per round of infection, μ, was fixed at $\mu = 5 \times 10^{-5}$. We corrected the 1.0×10^{-4} value given by Drake and Holland (9) in three ways, as
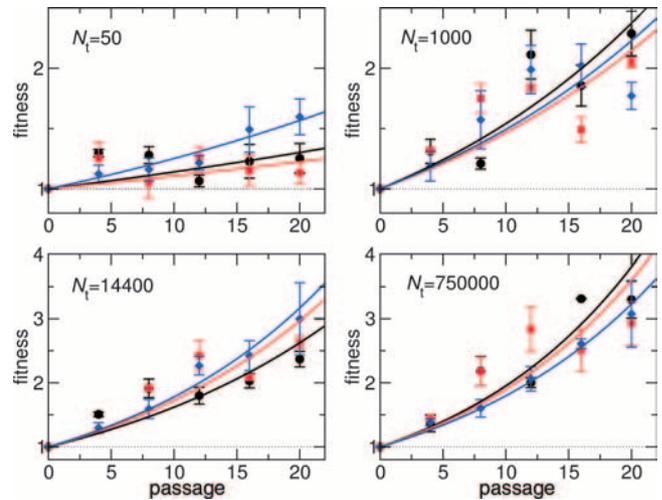

FIG. 1. Changes in relative fitness as a function of passage number for four representative N_t. Solid lines represent the least-square fit of the expression Vt (where t is the passage number) to the log-transformed fitness data. Larger populations experienced a more rapid fitness increase and less variation in fitness. Three different colors correspond to three replicate experiments. Vertical bars represent the standard deviations of the mean (standard errors) of fitness for each replica determined in the competition assays.

follows. (i) The mutation rate in the cited paper was given per RNA copy event. In the present work, we define the mutation rate per round of infection; one round of infection includes two RNA copy events: a negative-strand RNA is copied to positive-strand RNAs that are copied to negative-strand RNAs. (ii) The cited authors assumed four possible alleles per site (A, G, C, and T). In real data sets, only two alleles per site are usually found. (iii) Drake and Holland included insertions and deletions that, according to their estimate, increased the mutation rate by 40%. In the present work, we consider only point substitution. Combining these corrections, we obtain $\mu = (1.0 \times 10^{-4}) \times 2/(3 \times 1.4) = 5 \times 10^{-5}$. Using this value, we obtained the mutation rate for plaque-pooling passages as $\mu_P = 1.5 \times 10^{-4}$ and the mutation rate for dilution passages as $\mu_D = 1.0 \times 10^{-4}$.

To approximate the experimental data, we used four fitting parameters: the mean fitness effect per mutation (selection coefficient) s, the effective number of sites L, the mean number of sites in the initial population (MARM U) at which a mutation will be beneficial, which we denote as $k_{ref} = k_{av}(t = 0)$, and the variance of that number, d. Thus, a genome chosen randomly from the initial MARM U population will, on average, have k_{ref} sites at which a mutation leads to a fitness increase, and $L - k_{ref}$ sites at which a mutation leads to a fitness decrease.

Because we fit a relatively diverse data set and did not aim at high precision for fitting parameters that would exceed the accuracy of the model and the experiments, we employed manual fitting. The order in which different parameters were determined from different experiments was as follows. We determined the product sd for VSV MARM U from fitting the short-term adaptation rate (passages 0 to 20) at different values of N (see data in Fig. 1 to 3). The separate value of s (and, hence, d for MARM U) was determined by fitting the long-term time dependence of VSV MARM U, passages 20 to 50 (data in Fig. 3). We assumed that the value of s for all other strains of VSV was the same. We expressed the value k_{ref} in terms of s and $\mu_D L$ based on the maximum fitness measured in long-term dilution passages at population sizes on the order of 10^5, as given by $k_{ref} = (1/s)[\ln(10.5) + \mu_D L]$. The latter equation approximates the average long-term fitness with the steady-state value in a deterministic one-locus model, $\exp(-\mu_D L + sk_{ref})$, and the long-term fitness for MARM U (which we did not measure) with the value of 10.5 obtained for wild-type VSV (25).

To fit previous experiments of the evolution of several VSV strains between passages 0 and 20 (24 and new data in Fig. 4) and to estimate the effective number of evolving sites L (and thus, the effective genomic mutation rates $\mu_D L$ and $\mu_P L$), we assumed different values of d for three groups of strains with different evolutionary histories (see above). For strains MARM U and C, we determined d as explained in the previous paragraph. We adjusted the value of

d for MARM U35 and MARM X, as well as the value of *L*, to ensure no change of fitness between passages 0 and 20 at the critical N_t value and the correct rate of fitness decline due to Muller's ratchet at a smaller N_t value, as measured previously (24) and in this report. The value of *d* for MARM N, at the already known μL, was fit in the same way.

RESULTS

Measuring the speed of adaptation. We measured the speed of adaptation in MARM U VSV for N_t spanning more than 5 orders of magnitude, from an N_t of 4 to N_t of 7.5×10^5. For each N_t, we followed at least three replicate regimens during 20 passages. Fitness generally increased over time for all but the smallest N_t value. At an N_t of 4, 8, and 15, we observed that one out of five replicas declined in fitness, whereas the other four replicas increased in fitness. No fitness decrease occurred with any of the larger N_t. At small N_t, we found a significant amount of drift; fitness trajectories were erratic (not shown), and the total change in fitness over 20 passages varied substantially from replica to replica. By contrast, populations passaged at larger N_t values showed a consistent increase in fitness and less variability among replicas. In general, we found that fitness increased more in larger populations than in smaller populations. Typical fitness trajectories are shown in Fig. 1. Throughout this paper, we define the speed of adaptation, *V*, as the mean change in the logarithm of fitness per passage during the first 20 passages. We measured *V* for a given replica by calculating the natural logarithm for all fitness values in the replica and then fitting the function *Vt* (where *t* is time measured as the passage number) to the logarithm of the fitness values (Fig. 1).

Due to the effects of complementation after the first round of infection, the adaptation speed per passage is also the adaptation speed per first round of infection (see Materials and Methods). Note that the experimental protocols differed substantially between the replicas covering two intervals of transmission size N_t: from 4 to $3,000 \times 10^5$ and from 480 to 7.5×10^5. During plaque-pooling passages, there were three to four rounds of infection, and coinfection could take place only for mutants generated within each individual plaque. In contrast, during diluted passages, there were two rounds of infection, and the progeny from a cell infected during the first round could coinfect new cells with progeny from any other initially infected cell. However, our computational model predicted that the speed of adaptation over the initial 20 passages would be only marginally affected by these experimental differences, and this prediction was borne out by the experimental data (Fig. 2). In the regimen where *V* was measured using both methods, we found no detectable differences in the measured speed of adaptation between the two methods (two-sample *t* tests for differences in means found no significant differences, with *P* values of 0.52, 0.97, and 0.19 at N_t of 500, 1,000, and 3,000, respectively), providing further support for our claim that selection does not operate after the first round of infection in a passage.

We also used data from two other types of experiments for fitting. First, we used previously reported data on evolution of VSV strain MARM U on longer time scales, 50 passages, under dilution regimens (22). Second, we evolved the MARM N populations during 20 passages at its critical bottleneck size ($N_t = 2$), as well as the MARM U35 populations during 20
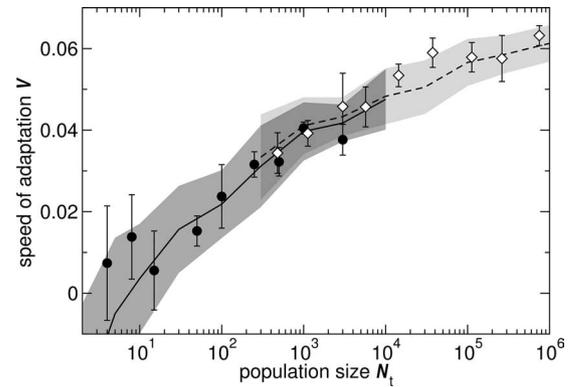

FIG. 2. Initial speed of adaptation, *V*, as a function of the transmission size, N_t. Symbols represent the measured adaptation speeds, which were averaged over the first 20 passages (Fig. 1), and vertical bars represent the associated standard deviations of the means. Black circles were obtained by plaque pooling, and open diamonds were obtained under dilution regimen conditions. Note the close match between the three values obtained by both methods. The solid line is the average adaptation speed predicted by the fitted simulation model for plaque-based passages, and the dashed line is the same value for dilution passages. The gray areas indicate 1 standard deviation around the predicted average adaptation speed. The average adaptation speed and standard deviation were calculated from 100 independent simulation runs. Fixed parameters are μ_P, 1.5×10^{-4} (for plaque pooling); μ_D, 1.0×10^{-4} (for dilutions). Fitting parameters are *s*, 0.39, k_{ref} = 6.25; *L* = 1,000; and *d* = 0.67.

passages at its critical bottleneck size ($N_t = 300$) and at a 10-fold smaller N_t ($N_t = 30$), all with plaque-pooling regimens. We combined these results with analogous published results for several VSV strains with different initial fitnesses, MARM X, MARM C, MARM U, and MARM N (24).

Fitting of the model. Our model has six parameters: population size (N_t), number of selected sites (*L*), number of deleterious alleles in the MARM U strain (k_{ref}), mutation rate per site per round of infection (μ), selection strength (*s*), and fitness diversity of the initial population (*d*). The population size was set by the experiment, and the mutation rate per round of infection was fixed at a μ value of 5×10^{-5} (leading to a mutation rate per passage of $\mu_P = 1.5 \times 10^{-4}$ for plaque-pooling regimens and $\mu_D = 1.0 \times 10^{-4}$ for dilution regimens). The remaining four parameters *s*, k_{ref}, *L*, and *d* had to be determined from fitting data, as described in Materials and Methods. We found good agreement between the simulation model and all data sets with a single set of the first three fitting parameters, *s* = 0.39, k_{ref} = 6.25, and *L* = 1,000, and three values of the fourth fitting parameter: *d* = 0.67 for VSV strains MARM U and MARM C, *d* = 0.29 for strains MARM U35 and MARM X, and *d* = 1.7 for strain MARM N (Fig. 2 to 4). Because the predicted dynamics are more sensitive to some parameters than to others, differently fitting parameters are determined with different levels of accuracy. For example, small changes in the values of *s* and *d* modify the predicted dynamics substantially, so the estimates for these values are relatively tight. In contrast, the least-critical parameter is the effective number of sites, *L*, and the value of 1,000 represents only a rough estimate. We determined *L* from the estimate of μ and the total deleterious mutation rate per genome, μL, which we adjusted to result in a simulated average fitness of a
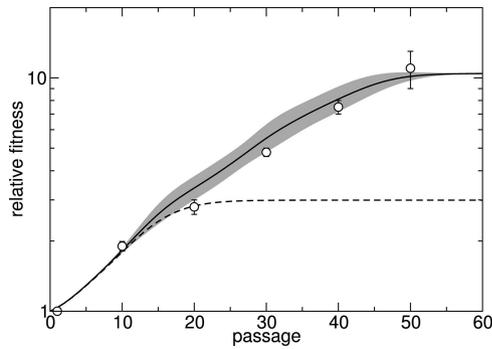
FIG. 3. Approach of virus fitness to mutation-selection balance for MARM U at a large N_t passaged under dilution regimens. Symbols represent MARM U fitness (22), and vertical bars represent the associated standard deviation of the mean. The solid line is a fit of the average fitness from the simulation model to the data at an N_t of 2×10^5. Other parameters are shown in the legend of Fig. 2. The gray area indicates 1 standard deviation around the average fitness. Average fitness and standard deviation were calculated from 100 independent simulation runs. The dashed line shows the model behavior if beneficial mutations are turned off. Thus, the initial 20 passages were dominated by the amplification of preexisting variants, whereas the later fitness increase was caused by the generation of new, beneficial mutations.

similar value at passage zero and passage 20 (Fig. 4). Changing the value of L to 800 or 1,200, with a small adjustment of other fitting parameters, produces a slightly less accurate fit (not shown).

The simulations faithfully reproduced the average speed of adaptation of MARM U over the first 20 passages as a function of N_t (Fig. 2). Figure 3 shows the approach of MARM U to mutation-selection balance over 50 passages, based on data previously reported (22). Because the total number of passages in this experiment did not suffice to observe the saturation of fitness, we assumed that the fitness of MARM U saturates at the same level as fitness of wt VSV studied previously on longer time scales (25). In Fig. 3 we have also indicated the prediction of our model in the absence of de novo generation

of beneficial mutations (dashed line). The fitness of the simulated viral population increases initially, as rare beneficial mutations are amplified, but levels off as soon as the fittest mutant present in the initial population has risen to fixation. We see that the initial 20 passages of the measured fitness trajectory are dominated by the selective amplification of preexisting mutations, whereas later fitness increases are due to newly generated beneficial mutations.

Figure 4 shows the predicted trajectories for five viral strains with different values of initial fitness (MARM U, MARM C, MARM U35, MARM X, and MARM N) at the critical bottleneck size, at which the average fitnesses at passage zero and at passage 20 are not significantly different (24 and the present report), as well as the experimental fitness trajectories of MARM N at its critical bottleneck size of 2. We also show two experimental cases, MARM X at an N_t of 5 (24), and MARM U35 at an N_t of 30, where we observed a fitness decrease (Fig. 4, left panel), as well as a case, MARM N at an N_t of 5, where there was an overall fitness increase (Fig. 4, right panel) (24). While none of the predicted curves simulated at the observed critical transmission size showed perfect fitness equality between passages zero and 20, the observed deviations are within the experimental uncertainty of the corresponding fitness measurements.

We obtained an estimate of the fitness of MARM U after 20 passages at a transmission size, N_t, of 1 from Table 1 in reference 23. We found an average fitness over 15 replicas of 0.65, with a standard deviation of the mean (standard error) of 0.09. At this N_t, we can solve our model analytically and find for the first and second moment of the fitness $\langle w \rangle = \exp[-Ut(1 - \exp(-s)]$ and $\langle w^2 \rangle = \exp[-Ut(1 - \exp(-2s)]$. From these expressions, we obtain a predicted mean fitness ± 1 standard deviation of 0.38 ± 0.23. This predicted value is significantly smaller than the measured value (two-sided t test, $P = 0.007$).

The values of the fitting parameters k_{ref} and L combined with the fixed parameter μ correspond to the effective mutation rate per genome per round of replication, $U = \mu L = 0.05$, and the effective rate of beneficial mutations per genome per
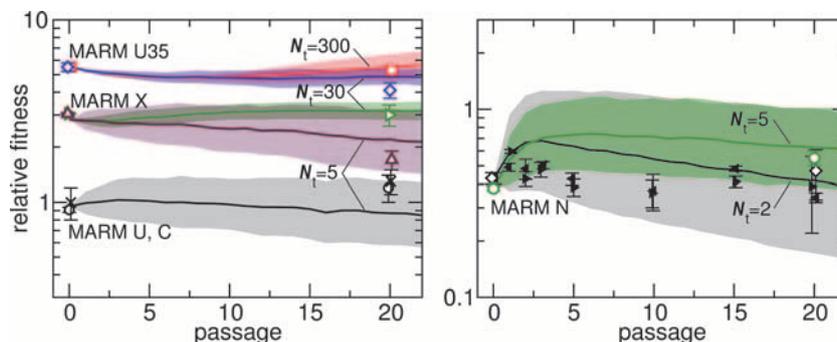

FIG. 4. Change in simulated fitness values close to the critical bottleneck size of N_t for several VSV strains passaged under plaque-pooling conditions. Open symbols represent measured fitness, averaged over six independent replicas, with error bars showing standard deviations of the means. Closed symbols in the right panel show fitness values at intermediate time points for two replicas at an N_t of 2. Data are both new and obtained from reference 24. For all symbols, we added a small amount of noise to the passage number to reduce symbol overlap. Solid lines and shaded areas indicate simulation results as described in the legends to Fig. 2 and 3. The observed fitness was approximately the same at passages 0 and 20 at the critical bottleneck size (24 and the present report). We also show simulation results for a subcritical transmission size for strains MARM X and MARM U35 and for a supercritical transmission size for strain MARM N. The transmission sizes for the different strains are indicated in the figure. The deleterious-allele-number variance d value is 0.67 for MARM U and MARM C, 0.29 for MARM X and MARM U35, and 1.7 for MARM N. Other parameter values are the same as those shown in Fig. 2 and 3.

round of replication is $U_b = \mu k_{ref} = 0.0003$ (these values have to be increased by a factor of 2 or 3 to apply to dilution or plaque-pooling regimens, respectively). The VSV genome is approximately 11 kb long, and thus, approximately $L/11$ kb = ~9% of all sites are under intermediate selection (have s values ≈0.39), whereas $k_{ref}/11$ kb = ~0.06% are sites of potentially beneficial mutations.

For MARM X, we also compared the standard deviation of fitness (given by the product sd) obtained from the fit to results from an independent experiment, in which the fitness of 98 subclones of that strain was measured (11). The predicted value of the standard deviation of fitness, $sd = 0.11$, agrees very well with the observed value of 0.097.

DISCUSSION

We have measured the initial speed of adaptation of a VSV strain, MARM U, over a range of population sizes spanning more than 5 orders of magnitude. We have found that we could explain this data set, as well as additional data sets obtained for several VSV strains (11, 22, 24, 25 and the present report), by a simple computational model of virus evolution, using a single set of fitting parameters for most of the data. The fitting parameters are all biologically relevant and encompass the variance in the number of mutations in the initial population, d, the mean effect of beneficial mutations, s, the number of sites in the genome that are under moderate selection, L, and the number of sites with potentially beneficial mutations, k_{ref}.

The model we used assumes a fixed fitness effect s for all mutations. Clearly, this assumption is an oversimplification; some mutations will have almost no effect on fitness, while others will be lethal. However, this assumption makes good modeling sense: we expect neutral or nearly neutral mutations to contribute substantially to sequence divergence, but by definition, they will not influence the fitness trajectory of a viral population. Therefore, we are justified in neglecting these mutations. Likewise, strongly deleterious or lethal mutations do not rise to large frequencies in the population. Hence, they do not affect the fitness trajectories either, and we can safely neglect them as well. We can also disregard mutations with large beneficial effects because they are unlikely in a strain, such as MARM U, that is well adapted to replication in BHK-21 cells. Thus, only mutations with a moderate effect will alter the trajectories, and we describe these mutations by their mean (or typical) effect, s, determined by the time scale of several dozens of generations studied in an experiment. However, since these intermediate mutations cannot occur at all sites in the viral genome, we need a second parameter, L, which represents the number of sites at which these mutations can occur. We define "almost neutral" sites (not included in L) as those that can experience beneficial or deleterious mutations with selection coefficients of s that are much less than 0.4; sites with an s of approximately 1 or larger are considered strongly deleterious.

Why use such a simplified model? Including all the existing evolutionary factors into a model is not possible, as such a model would become overly complex and contain more fitting parameter than can be determined from data. The model complexity is dictated by the aim of the study and the size of the data set. Whether a certain factor is important quantitatively for a data set is never obvious a priori and has to be determined by the ability of a simpler model to fit data.

In particular, there is little doubt that the value of s is broadly distributed among sites, but only the characteristic values of s corresponding to the time scale of evolution studied in the present work are relevant for our model. For larger time scales, the characteristic value of s would be smaller, and the number of sites, L, larger. Because our simple model fits data well and because parameters of a more complex model could not be determined from our data set, we are not only allowed but also required to make the simplification of constant s.

We also neglect epistasis in our model. Once again, we made this modeling decision not because we believe that there is no epistasis in VSV. To the contrary, numerous studies show that epistatic interactions are widespread in RNA viruses, including VSV (1–3, 21, 37). Our assumption is that for the initial phase of adaptation we study here, whatever epistatic interactions may exist in VSV are irrelevant on average and can be subsumed into a single effective selection coefficient s in an independent-sites model. The good agreement between the data and the fitted model justifies this approach. To produce a meaningful fit of a model with epistasis, we would need a much more extensive data set, including measurements such as those presented in Fig. 2 for multiple virus strains and possibly for larger time scales.

We would like to emphasize that our work does not aim to determine the rate of occurrence and distribution of effects of new mutations, unlike mutagenesis or mutation-accumulation experiments (see reference 12 for a recent review). Our work identifies instead the phenotypically relevant standing variation in a viral population. Also, our work differs from statistical analyses of mutation-accumulation experiments in that we follow a first-principle modeling approach. The goal of a statistical analysis is to extract information from noise, using relatively simple mathematical expressions that do not, usually, correspond to mechanistic models of the biological system. The goal of first-principle modeling is to find which factors in a mechanistic model are most important and to estimate the model parameters that describe these factors. To give a concrete example, in the statistical approach, one can simply assume that beneficial and deleterious mutations have different effect sizes and estimate the effect sizes independently. In the first-principle modeling approach, on the other hand, beneficial and deleterious mutations are intrinsically linked, just two sides of the same coin, where every fixed beneficial mutation is a potential deleterious mutation and vice versa. We are not arguing here that one approach is inherently better than the other. We see them as complementary. The main strength of the first-principle modeling approach is that it leads to a mechanistic understanding of virus evolution.

In our model, we assume that selection operates only in the first round of infection in each passage, because later rounds of infection take place at high MOI, where complementation masks fitness differences among mutants (26). This assumption would be problematic if superinfection exclusion were keeping the mean number of infecting genomes per cell low, even at nominally high MOI. However, in our experiments, superinfection exclusion could only have limited effects in restricting the number of PFU per cell that initiated the next round of infection. The different mechanisms of superinfection exclu-

sion operating in VSV-infected cells become effective only 2 to 4 h postinfection and are relatively inefficient. At 2 h postinfection, cells can still be superinfected with 75% of the available virus (34). Even if 90% of the progeny after one round of infection were unable to infect the new cells due to superinfection exclusion or because the virus does not encounter a new cell and remains in the medium, the MOI in the second round of infection in our experiments would still be 25 PFU/cell. Finally, we have demonstrated that restrictions to coinfection are not sufficient to limit complementation and that the fitness of MARM N during a standard competition, which includes two rounds of infection under the same conditions as in our diluted regimens, can be explained only if we assume that complementation operates during the second round of infection at high MOI (40).

We found here an s value of 0.39 and an L value of $\approx 1,000$ for a genome that is approximately 11 kb long. According to our model, approximately 9% of the genome is under moderate selection, and the remaining 91% is either under strong purifying selection or is selectively neutral. As explained above, our fitting procedure cannot distinguish between these two cases. We subdivided the 9% of the genome under moderate selection into sites that are under negative selection, i.e., sites at which mutations will lead to a fitness decrease and sites that are under positive selection, i.e., sites at which mutations will lead to a fitness increase. We achieved this subdivision for MARM U with the fitting parameter k_{ref}, which counts the mean number of positively selected sites. For other strains, the number of positively selected sites correlated inversely with their relative fitness. For MARM U, we found a fairly low number of beneficial sites, $k_{ref} = 6.25$ sites per genome. This value is consistent with complete-genome sequencing studies of evolved wt VSV: at the time that the wt reached the mutation-selection balance during adaptation to BHK-21 cells, the consensus sequences of the evolved genomes from different replicas accumulated between two and eight substitutions, with an average of six substitutions per genome (27).

We used additional data for the change of fitness between passages 0 and 20 from several other VSV strains (24 and this report) to obtain an additional test of the model and to estimate the value of L as stated above. Initially, we tried to use the value of d obtained for strain MARM U for all other strains, but we could not obtain an acceptable fit (results not shown). Therefore, we grouped the various strains into three groups and used a separate d value for each group. The first group included strains that were derived from less than five amplification passages after the isolation of a single clone (MARM U and MARM C); the second group included strains that were derived from 35 or more large-population passages (MARM U35 and MARM X); a third value of d was used for strain MARM N. We obtained an acceptable fit when we assumed that the first group had a larger d ($d = 0.67$) than the second group ($d = 0.29$). At first glance, this result is surprising, because the strains that have experienced more passaging have had more time to accumulate mutations and have a higher within-population diversity (unpublished data). However, this result makes sense if we take into account the fact that d measures not the within-population variation, as characterized by the genetic distance or the average number of differences from the current consensus sequence, but the variation in the number of deleterious alleles per genome (i.e., the fitness variation). For example, if all the genomes have exactly the same number of deleterious alleles located at different sites, then they all have the same fitness; thus, d is equal to zero, but the genetic distance is non-zero and can be substantial. During limited amplification of a single virion, mutant virus will be produced, but there will not be enough time for negative selection to eliminate deleterious mutants. By contrast, for a strain that has experienced extensive adaptation, selection will effectively remove a large fraction of genomes with low fitness, leading to a smaller value of d. Also, d is related only to sites that are under intermediate selective pressure, whereas the total genetic distance also includes neutral or nearly neutral mutations. The consistency between sd values predicted with the model (0.11) and those obtained experimentally (0.097) for MARM X (11) further supports our choice of different parameters for the three groups of strains: (i) MARM X and MARM U35, (ii) MARM U and MARM C, and (iii) MARM N. However, while we chose to maintain s at a fixed value and use variable values for d, the results of the simulations for the strains other than MARM U (i.e., Fig. 4) would be identical if we maintained a fixed value of d and varied the value of s. None of the available data allows us to distinguish between the two alternatives, and it may be possible that both parameters change to some degree among strains.

The experimental data at a very small N_t value had a tendency to deviate from the averages predicted from the simulations more than the data obtained at larger N_t values (see MARM N data in Fig. 4 and the data for MARM U at $N_t = 1$ [not graphed]), even though the experimental results generally fell within 1 standard deviation from the predicted average. Our model is based on two assumptions that may not be valid at very small N_t values. First, we assume that the effective population size is given by the value N_t. Since the virus population expands during a passage, this assumption always underestimates the true effective population size. For sufficiently large N_t, a small error in our estimate of the effective population size does not significantly alter the predicted viral dynamics. However, at very small N_t values, the simulation becomes sensitive to small differences in the population size. For example, the predicted mean fitness after 20 passages for MARM U is almost twice as large at an N_t of 2 as at an N_t of 1. Thus, a small error in the estimated effective population size can easily lead to significant discrepancies between the simulation and the measurement at small N_t values. Second, we assume that selection operates only during the first round of infection in a passage but ceases to have any effect in later rounds of infection, when coinfection is frequent. If there is some residual selection in later rounds of infection, then this selection will also counteract the loss of fitness at very small N_t values and thus increase the discrepancy between experiment and simulation. A third potential factor to consider in the case of plaque-to-plaque regimens is sampling bias. For counting plaques, we stain the monolayers with crystal violet, and plaques appear as clear or turbid areas. However, this treatment kills both cells and virus, so for plaque picking, the monolayers cannot be stained. Plaques that are very small, most of which would represent deleterious mutants, may be easy to miss and be left out during sampling, and the result would be a slower-than-expected decline in fitness.

For the mutation rate, we used a value obtained from reference 9 and subjected to various corrections to make it applicable to our experiments. Conceivably, either the original value or the corrections may be not accurate. Fortunately, we found that our results are weakly dependent on the specific mutation rate we use. For example, let us assume that the mutation rate value of μ is 4×10^{-4}, i.e., 2.5- to 4-fold higher than the values 1×10^{-4} and 1.5×10^{-4} that we used for the two types of experimental regimens. As we have checked (simulation results not shown), the best-fit value of s changes from 0.39 to 0.28, that of d (for MARM U and MARM C) from 0.69 to 1.0, that of k_{ref} from 6 to 9, and that of L from 1,000 to 500. Thus, only the effective number of sites changes by a factor of 2 and should be regarded as an estimate; all other parameters change less. Because it is unlikely that the error in the mutation rate is as large as fourfold, the actual error in our parameters is even smaller, especially since for MARM X the d value of 0.29 is in excellent agreement with direct measurements of its fitness distribution (11).

One data set used in our fit was the fitness trajectory of MARM U's approach to mutation-selection balance (Fig. 3). In the particular data set we used, fitness had been measured only up to passage 50 and had not leveled off at this passage. Therefore, for this data set, we cannot be sure that fitness would have leveled off at exactly the time point and fitness value we assumed. Nevertheless, we believe that our assumptions are reasonable, because MARM U differs by only a single neutral marker from the VSV wild type, and the approach to mutation-selection balance had previously been measured more extensively in the wild type (25). If we assumed that MARM U would reach a somewhat higher equilibrium fitness level (e.g., 13 instead of 10.5), we found that fitting the parameter k_{ref} would change slightly, but our general conclusions would not change (data not shown).

The early fitness trajectory, encompassing approximately the first 20 passages, was dominated by the amplification of preexisting variation, whereas later adaptation was caused by the de novo generation of beneficial mutations (Fig. 2). This result confirmed an earlier prediction for the same model by Tsimring et al. (35), who argued that viral adaptation proceeds in a two-phase process, where during the first phase, the highest-fitness minority mutants grow to a macroscopic proportion in the population, and then in the second phase, the population as a whole moves toward higher fitness (33).

To summarize, we have shown that the initial adaptation of VSV populations is well described by a standard, multiplicative, multisite model and that by fitting the model to several datasets, we can obtain estimates of the amount of preexisting variation in a VSV population, the mean selection coefficient, the beneficial mutation rate, and the number of sites under selection. The estimate of the amount of preexisting variation we obtained agrees well with an independent measurement of the same quantity, which lends additional support to our approach. Our method can be generalized and applied to other rapidly evolving viruses, including those with recombination (for instance, HIV). The method can also be extended to DNA-based microorganisms, such as yeast, which can reproduce either asexually (6) or sexually (15). DNA microbes have slower evolutionary rates, and the values of parameters such as s and μL will be smaller, but our approach can still be used

to describe the transitional period from a genetically diverse population to a stationary regimen or to mutation-selection balance.

REFERENCES

1. Ayme, V., J. Petit-Pierre, S. Souche, A. Palloix, and B. Moury. 2007. Molecular dissection of the potato virus Y VPg virulence factor reveals complex adaptations to the pvr2 resistance allelic series in pepper. J. Gen. Virol. 88:1594–1601.
2. Bonhoeffer, S., C. Chappey, N. T. Parkin, J. M. Whitcomb, and C. J. Petropoulos. 2004. Evidence for positive epistasis in HIV-1. Science 306:1547–1550.
3. Burch, C. L., P. E. Turner, and K. A. Hanley. 2003. Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design. J. Evol. Biol. 16:1223–1235.
4. Coffin, J. M. 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science 267:483–489.
5. Cohen, E., D. A. Kessler, and H. Levine. 2005. Recombination dramatically speeds up evolution of finite populations. Phys. Rev. Lett. 94:098102.
6. Desai, M. M., D. S. Fisher, and A. W. Murray. 2007. The speed of evolution and maintenance of variation in asexual populations. Curr. Biol. 17:385–394.
7. Domingo, E., C. Escarmis, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. S. Novella, and J. J. Holland. 1996. Basic concepts in RNA virus evolution. FASEB J. 10:859–864.
8. Domingo, E., and J. J. Holland. 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. 51:151–178.
9. Drake, J. W., and J. J. Holland. 1999. Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. USA 96:13910–13913.
10. Duarte, E., D. Clarke, A. Moya, E. Domingo, and J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc. Natl. Acad. Sci. USA 89:6015–6019.
11. Duarte, E. A., I. S. Novella, S. Ledesma, D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland. 1994. Subclonal components of consensus fitness in an RNA virus clone. J. Virol. 68:4295–4301.
12. Eyre-Walker, A., and P. D. Keightley. 2007. The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8:610–618.
13. Gerrish, P. J., A. Colato, A. S. Perelson, and P. D. Sniegowski. 2007. Complete genetic linkage can subvert natural selection. Proc. Natl. Acad. Sci. USA 104:6266–6271.
14. Gerrish, P. J., and R. E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. Genetica 102–103:127–144.
15. Goddard, M. R., H. C. Godfray, and A. Burt. 2005. Sex increases the efficacy of natural selection in experimental yeast populations. Nature 434:636–640.
16. Holland, J. J., J. C. de la Torre, D. K. Clarke, and E. Duarte. 1991. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. J. Virol. 65:2960–2967.
17. Johnson, J. A., K. K. Rompay, E. Delwart, and W. Heneine. 2006. A rapid and sensitive real-time PCR assay for the K65R drug resistance mutation in SIV reverse transcriptase. AIDS Res. Hum. Retrovir. 22:912–916.
18. Kessler, D. A., H. Levine, D. Ridgway, and L. Tsimring. 1997. Evolution on a smooth landscape. J. Stat. Physics 87:519–544.
19. Lefrancois, L., and D. S. Lyles. 1982. The interaction of antibody with the major surface glycoprotein of vesicular stomatitis virus. I. Analysis of neutralizing epitopes with monoclonal antibodies. Virology 121:157–167.
20. Nájera, I., A. Holguín, M. E. Quiñones-Mateu, M. A. Muñoz-Fernández, R. Nájera, C. López-Galíndez, and E. Domingo. 1995. pol gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. J. Virol. 69:23–31.
21. Novella, I. S. 2004. Negative effect of genetic bottlenecks on the adaptability of vesicular stomatitis virus. J. Mol. Biol. 336:61–67.
22. Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland. 1995. Exponential increases of RNA virus fitness during large population transmissions. Proc. Natl. Acad. Sci. USA 92:5841–5844.
23. Novella, I. S., and B. E. Ebendick-Corpus. 2004. Molecular basis of fitness loss and fitness recovery in vesicular stomatitis virus. J. Mol. Biol. 342:1423–1430.
24. Novella, I. S., S. F. Elena, A. Moya, E. Domingo, and J. J. Holland. 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. J. Virol. 69:2869–2872.
25. Novella, I. S., J. Quer, E. Domingo, and J. J. Holland. 1999. Exponential fitness gains of RNA virus populations are limited by bottleneck effects. J. Virol. 73:1668–1671.
26. Novella, I. S., D. D. Reissig, and C. O. Wilke. 2004. Density-dependent selection in vesicular stomatitis virus. J. Virol. 78:5799–5804.

27. **Novella, I. S., S. Zarate, D. Metzgar, and B. E. Ebendick-Corpus.** 2004. Positive selection of synonymous mutations in vesicular stomatitis virus. J. Mol. Biol. **342:**1415–1421.
28. **Orr, H. A.** 2000. The rate of adaptation in asexuals. Genetics **155:**961–968.
29. **Pringle, C. R., and A. J. Easton.** 1997. Monopartite negative-strand RNA genomes. Semin. Virol. **8:**49–57.
30. **Richman, D. D., D. Havlir, J. Corbeil, D. Looney, C. Ignacio, S. A. Spector, J. Sullivan, S. Cheeseman, K. Barringer, D. Pauletti, et al.** 1994. Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy. J. Virol. **68:**1660–1666.
31. **Rouzine, I. M., and J. M. Coffin.** 2005. Evolution of human immunodeficiency virus under selection and weak recombination. Genetics **170:**7–18.
32. **Rouzine, I. M., A. Rodrigo, and J. M. Coffin.** 2001. Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. Microbiol. Mol. Biol. Rev. **65:**151–185.
33. **Rouzine, I. M., J. Wakeley, and J. M. Coffin.** 2003. The solitary wave of asexual evolution. Proc. Natl. Acad. Sci. USA **100:**587–592.
34. **Simon, K. O., J. J. Cardamone, Jr., P. A. Whitaker-Dowling, J. S. Youngner,** and C. C. Widnell. 1990. Cellular mechanisms in the superinfection exclusion of vesicular stomatitis virus. Virology **177:**375–379.
35. **Tsimring, L. S., H. Levine, and D. A. Kessler.** 1996. RNA virus evolution via a fitness-space model. Phys. Rev. Lett. **76:**4440–4443.
36. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, et al.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature **373:**117–122.
37. **Weinreich, D. M., R. A. Watson, and L. Chao.** 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59:**1165–1174.
38. **Wilke, C. O.** 2004. The speed of adaptation in large asexual populations. Genetics **167:**2045–2053.
39. **Wilke, C. O., and I. S. Novella.** 2003. Phenotypic mixing and hiding may contribute to memory in viral quasispecies. BMC Microbiol. **3:**11.
40. **Wilke, C. O., D. D. Reissig, and I. S. Novella.** 2004. Replication at periodically changing multiplicity of infection promotes stable coexistence of competing viral populations. Evolution **58:**900–905.