

The cost of replication fidelity in an RNA virus

Victoria Furió, Andrés Moya, and Rafael Sanjuán*

Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, P.O. Box 22085, 46071 València, Spain

Edited by Tomoko Ohta, National Institute of Genetics, Mishima, Japan, and approved May 31, 2005 (received for review February 8, 2005)

It is often argued that high mutation rates are advantageous for RNA viruses, because they confer elevated rates of adaptation. However, there is no direct evidence showing a positive correlation between mutation and adaptation rates among RNA viruses. Moreover, theoretical work does not argue in favor of this prediction. We used a series of vesicular stomatitis virus clones harboring single amino acid substitutions in the RNA polymerase to demonstrate that changes inducing enhanced fidelity paid a fitness cost, but that there was no positive correlation between mutation and adaptation rates. We demonstrate that the observed mutation rate in vesicular stomatitis virus can be explained by a trade-off between replication rate and replication fidelity.

fitness trade-off | mutation rate | adaptation | experimental evolution

Mutation provides the allelic variation that natural selection can act on. Conversely, mutation rates show phenotypic variability, which turns out to be a target for selection. RNA viruses are characterized by high mutation rates compared with most DNA systems (1), due mainly to the lack of exonuclease proofreading activity displayed by their RNA polymerases (2). Current evidence suggests that this high mutation rate cannot simply be attributed to biochemical restrictions: in HIV-1, several antimutator retrotranscriptases have been described (3–6); Pfeiffer and Kierkegaard (7) showed that a single-residue mutation in the RNA polymerase gene of type 1 poliovirus can confer resistance to ribavirin through a 3-fold increase in fidelity; Pugachev *et al.* (8) suggested that mutation rate in the yellow fever virus could be as low as 2×10^{-7} mutations per nucleotidic site per replication round. Such evidence demonstrates the need for an evolutionary model that accounts for high mutation rates in RNA viruses beyond a purely mechanistic level.

It is often argued that high mutation rates in RNA viruses are favored, because they confer a greater adaptive capacity (9, 10). However, there is still no experimental proof for this selective advantage: no evidence has been found to support a positive correlation between mutation and adaptation rates, as one would expect according to this hypothesis. A substantial part of our knowledge on the nexus between mutation and adaptation comes from experiments exploring the dynamics of mutator and antimutator genotypes in *Escherichia coli* (e.g., ref. 11), where high mutation rates have been associated with increased population fitness, because of the genetic hitchhiking of the mutator allele with beneficial changes produced at other loci. The hitchhiking hypothesis (12) might explain why high error rates could have risen and could have been maintained in RNA viruses, especially in the absence of recombination. On the other hand, the vast majority of mutations having a phenotypic effect are deleterious. Then, short-term selection should favor lower mutation rates, to minimize the genetic load (Fig. 1). Moreover, under high mutation rates, favorable alleles will frequently arise in genomes carrying at least one deleterious mutation, thus reducing their fixation probability. Taking these factors into account, increased mutation rates should not necessarily confer faster adaptation rates, but rather the optimal evolutionary solution should be reached at intermediate mutation rates (13, 14).

Alternatively, selective pressure favoring low mutation rates to minimize the genetic load could be counterbalanced by selection in the opposite sense, because of a biochemical cost of replication fidelity (15–17). This being true, a positive correlation between mutation rate and fitness should be observed (Fig. 1). Here, using a series of single-nucleotide vesicular stomatitis virus (VSV) polymerase mutants, we report evidence supporting the latter possibility. Our results suggest an evolutionary explanation for high mutation rates in RNA viruses.

Methods

Virus. VSV is a lytic negative-stranded RNA virus from the *Rhabdoviridae* family. Its infection cycle begins with viral adsorption and entry, followed by the synthesis of intermediate positive-stranded molecules, which act as templates for transcription and for the synthesis of negative-stranded molecules. The latter are then packaged and released into the medium. However, at any given time in the course of a cell culture infection, the majority of particles are located intracellularly, because virions accumulate in the cytoplasm before lysis. Negative-stranded RNA viruses have very low levels of recombination, and hence they are effectively asexual (18).

Site-Directed Mutagenesis. A full-length infectious cDNA clone was used for creating single-substitution mutants in the L polymerase gene. The amino acid replacements introduced were I270V, I131N, F910S, G1348A, L549F, K2054E, I2003V, and N85K. As described (19), change N85K was chosen at random, whereas the remaining reproduced polymorphisms occurring spontaneously in laboratory populations (change I270V has not been reported previously). Site-directed mutagenesis reactions were carried out by using *Pfu* DNA polymerase (Promega). Products were digested with DpnI (Stratagene) to remove the parental strands and transformed into ultracompetent *E. coli* cells (Stratagene). Sequencing of the cDNAs was performed to confirm that each desired mutation was incorporated.

Virus Recovery from cDNA Clones. Baby hamster kidney (BHK₂₁) cells (American Type Culture Collection) were infected with a recombinant vaccinia virus expressing T7 RNA polymerase. After incubation, cells were cotransfected with the mutant cDNA clone plus three support plasmids encoding P, L, and N proteins, using Lipofectamine and Plus Reagents (Invitrogen) (19). The supernatant was harvested 96 h postinoculation (hpi), and residual vaccinia was removed by filtering the supernatant through 0.2- μ m membranes (Millipore). Posttransfection titers, which ranged from 10^4 to 10^6 plaque-forming units (pfu)/ml, were equalized to $\approx 5 \times 10^6$ pfu/ml before fluctuation tests and competition assays (19).

Luria–Delbrück Fluctuation Tests (20). A 96-well plate containing $\approx 10^4$ BHK₂₁ cells per well was infected with ≈ 500 pfu per well

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: hpi, hours postinoculation; pfu, plaque-forming units; $ss\ l^{-1}r^{-1}$, substitutions per locus per replication cycle; VSV, vesicular stomatitis virus; BHK₂₁, baby hamster kidney; ML, maximum likelihood; CI, confidence interval.

*To whom correspondence should be addressed. E-mail: rafsafer@ibmcp.upv.es.

© 2005 by The National Academy of Sciences of the USA

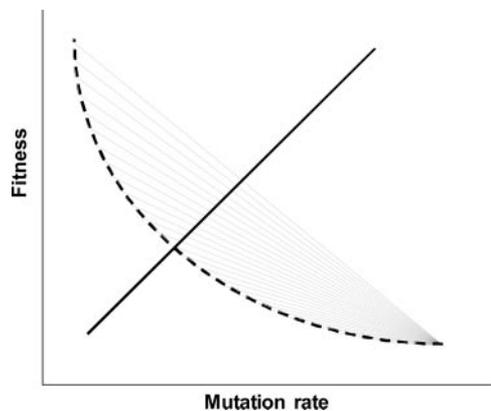


Fig. 1. Expected relationships between fitness and mutation rate. Without a direct effect of the mutation rate modifier allele on fitness, the mean fitness is expected to exponentially decrease as a function of mutation rate because of increasing genetic load (dashed lines). By contrast, if there is a cost of fidelity, fitness can rise as the mutation rate increases. The solid line has only qualitative purposes.

and then incubated until $\approx 3 \times 10^4$ pfu were expected. During this time, no more than a single infection cycle elapsed on average, thus minimizing the effects of selection. Before plating, cultures were freeze-thawed to allow the liberation of intracellular particles. We used 84 wells for resistance screening and 12 wells for titration. As a phenotypic marker, we used resistance to monoclonal antibody I₁, which is conferred by each of the following single-amino acid (and single-nucleotide) substitutions in the glycoprotein gene: D257G, D259N, D259A, D257V, D257Y, and A263E (21). The presence of resistant genotypes in each well was monitored by plating all of the undiluted supernatant (100 μ l) into 60-mm plates containing $\approx 10^6$ BHK₂₁ cells and I₁ at a concentration that completely inhibits wild-type growth. This procedure largely prevents the complications of phenotypic masking (22). After 45-min incubation to allow viral adsorption, medium supplemented with I₁ 25% vol/vol plus agar 0.4% (Sigma) was added, and cultures were incubated for 28 hpi. Then, lysis plaques were visualized by staining the cell monolayer with a solution of crystal violet 2% (Sigma) in formaldehyde 10% (Panreac, Spain). Usually plaques are visible at 20 hpi, but staining was done after 8 additional hours to ensure that no resistant plaque was missed. Despite the freeze-thawing, there is no guarantee that all of the produced viruses were indeed plated. However, this would not affect mutation rate estimation, provided that the screened nucleotide substitutions are selectively neutral. Further, analyses were done by using relative mutation rates, hence preventing this putative problem from affecting the results. Titration was performed by plating 100 μ l of a 300-fold diluted supernatant. Plating conditions were the same as for mutant screening, except that no I₁ was added, and staining was done at 18–22 hpi. For ancestral genotypes, measures were performed in three replicates and six replicates for the wild type. Because measures for the evolved genotypes were done separately in time, six additional replicates were performed for the wild type. For this reference genotype, the null-class mutation rate estimation (see below) was $(1.23 \pm 0.30) \times 10^{-5} \text{ ss} \cdot \text{l}^{-1} \cdot \text{r}^{-1}$ in the first block and $(1.29 \pm 0.22) \times 10^{-5} \text{ ss} \cdot \text{l}^{-1} \cdot \text{r}^{-1}$ (ss l⁻¹·r⁻¹, substitutions per locus per replication cycle) in the second block. No differences between these two blocks were observed (*t* test, $n = 12$, $P = 0.879$), indicating that the results of the fluctuation tests were reproducible. To estimate relative mutation rates, each value was divided per wild-type mean belonging to the same block.

Discarding Previously Existing Resistant Genotypes. During fluctuation tests, for the sake of reproducibility, each well was seeded with ≈ 500 pfu, and therefore it is possible that I₁ resistant genotypes were already present in the initial mixture. Using the values obtained from the 12 wells reserved for titration, we calculated the number of new viruses generated by each initial particle, Q , as $Q = N_f / N_i$, where N_f is the final number of pfu determined from titrations, and $N_i = 500$ pfu. If, in a given well, resistant genotypes were already present at the beginning of the infection, then at least Q resistant pfu should be observed in the corresponding 60-mm plate. This hypothesis can be tested by using a one-tail one-sample *t* test, in which the 12 Q values are compared with the observed number of resistant pfu. Significance thresholds were adjusted according to the Bonferroni correction. In the case of the ancestral genotypes, of the 2,520 wells used for resistance screening, the possibility of previously existing resistant genotypes could not be rejected in six cases, whereas in the evolved populations, this occurred in 10 of 2,520 cases. The latter were discarded for further analyses, although their inclusion did not affect the results.

Mutation Rate Estimation. Mutation rates per locus per replication cycle were estimated by using five different methods. On the one hand, if viruses replicated in a strictly binary mode, we could estimate the mutation rate as $\mu_{b-f} = 2f / \log_2(N_f / N_i)$, where f is the frequency of mutants (23), or by fitting the observed distribution of mutants to a Luria–Delbrück distribution by using a maximum likelihood (ML) algorithm (11). These values will be referred to as μ_{b-ML} and were calculated by using the FLUCMXML program, kindly provided by P. J. Gerrish (Mexican Institute of Petroleum, Mexico City). On the other hand, if replication was fully linear, mutation rates could be estimated as $\mu_{l-f} = f/c$ (24), where c is the number of infection cycles elapsed between N_i and N_f (25), or by fitting the data to a Poisson distribution (μ_{l-ML}), where μ_{l-ML} maximizes $\log L = \sum_n P(n|\lambda)$, P being the Poisson probability density function, $\lambda = \mu(N_f - N_i)$, and n the number of mutants per 96 wells. The variance of μ_{l-ML} was calculated as

$$\sigma_\mu^2 = \sum_n \frac{1}{\sqrt{nE(\delta \log P_n / \delta \mu)^2}} \quad (26).$$

Finally, the mutation rate can be estimated using the null-class method, such that $\mu_0 = -\log P_0 / (N_f - N_i)$, where P_0 is the proportion of plates with no visible growth (20).

Relative Fitness Assays. We assessed the fitness of each mutant relative to a reference strain [monoclonal antibody-resistant mutant (MARM) Rafael Sanjuan Verdeguer (RSV)], which is resistant to mAb I₁ (19). To do so, we seeded $\approx 5 \times 10^3$ pfu of each genotype into $\approx 10^5$ cells in 24-well plates. Titers of both genotypes were determined by plating in the presence and absence of I₁. Intrinsic growth rates (r) were calculated as the slope of the log-titer against time (hpi) during the period of exponential growth. Fitness assays were done in triplicate, and relative fitness was calculated as $W = e^{r_{\text{mutant}} - r_{\text{MARM RSV}}}$. Preliminary competition assays showed that the wild type is effectively neutral relative to the MARM RSV ($W = 1.003 \pm 0.009$; $t_{59} = 0.405$, $P = 0.687$).

Serial Infection Passages. An inoculate containing $\approx 5 \times 10^3$ pfu was added to $\approx 10^5$ BHK₂₁ cells in 24-well plates. After 24 hpi, supernatants were harvested and titrated (titers typically reached 10^9 pfu/ml). Then, fresh cells were used to repeat the infection until 10 passages were completed. All serial infection passages were done in triplicate. Adaptation rates can be calculated as $R_W = W_f / W_i$, where W_f and W_i stand, respectively, for the evolved and ancestral fitness relative to the wild type.

Table 1. Observed mutant distribution in each of the six Luria–Delbrück tests for the wild type

Number of mutants	Fluctuation test					
	1	2	3	4	5	6
0	31	40	50	32	39	39
1	22	22	19	23	28	19
2	19	12	9	17	9	17
3	5	6	3	5	2	2
4	3	0	2	3	1	2
5	2	1	0	2	0	2
6	1	0	0	1	1	0
>6	1	3	1	1	4	3
P	0.173	<0.001	0.020	0.104	<0.001	<0.001

The last row give the probability of the Poisson model according to the Kolmogorov–Smirnov test.

Results

Mutation Rate Estimation. We performed six Luria–Delbrück fluctuation tests for the wild type, and we estimated the mutation rate of VSV with different available methods. When binary replication was assumed, we obtained $\mu_{b-f} = 6.88 \times 10^{-6}$ [95% confidence interval (CI₉₅): $3.36 \times 10^{-6} - 1.04 \times 10^{-5}$] ss l⁻¹·r⁻¹ using the observed mutant frequency, and $\mu_{b-ML} = 8.95 \times 10^{-6}$ (CI₉₅: $7.97 \times 10^{-6} - 1.00 \times 10^{-5}$) ss l⁻¹·r⁻¹, fitting the observed distribution of the number of mutants per well to the Luria–Delbrück distribution. However, VSV is unlikely to replicate in a binary manner, because *in vivo* studies have revealed that the number of negative-strand progeny genomes exceeds that of positive-strand intermediate genomes at least 6-fold (27).

Assuming a linear replication mode, we obtained $\mu_{l-f} = 3.54 \times 10^{-5}$ (CI₉₅: $1.73 \times 10^{-5} - 5.34 \times 10^{-5}$) ss l⁻¹·r⁻¹ using the observed mutant frequency, and $\mu_{l-ML} = 2.19 \times 10^{-5}$ (CI₉₅: $0 - 3.35 \times 10^{-4}$) ss l⁻¹·r⁻¹, fitting the observed distribution of the number of mutants per well to a Poisson distribution. To assess whether the linear replication assumption is acceptable, we tested the fit of the data to the expected mutant distribution, i.e., the Poisson distribution (28). A Kolmogorov–Smirnov test (Table 1) indicated that the distribution is unlikely to be Poisson, and hence the VSV mode of replication cannot be assumed to be purely linear.

Finally, using the null-class method, we obtained $\mu_0 = 1.23 \times 10^{-5}$ (CI₉₅: $4.68 \times 10^{-6} - 2.00 \times 10^{-5}$) substitutions per locus per replication event. This latter method is preferable, because it allows obviating considerations about the replication mode, and it also has the advantage of being unaffected by a putative lack of neutrality of the phenotypic marker. For these reasons, we will use μ_0 to estimate the mutation rate hereafter. First, note that the mutation rate is not expressed per replication cycle but rather per replication event, because no mechanism is assumed. For the two previous approaches (binary and linear), both unities are equivalent. Second, the null-class method gives intermediate mutation rate estimations, as expected if the replication mode was a compound between binary and linear replication.

Correlation Between Fitness and Mutation Rate. First, mutation rates differed slightly but significantly (one-way ANOVA, $P = 0.007$) among the eight clones harboring a single amino acid replacement in the *L* (RNA polymerase) gene. The relatively high error of null-class estimations did not allow us to find significant differences between each single clone and the wild type, but according to μ_{b-ML} CI₉₅ (which are much smaller than those of μ_0), substitutions I131N and I270V produced antimutator phenotypes relative to the wild type, whereas mutator phenotypes were not observed. The strong correlation between μ_0 and μ_{b-ML} (Pearson correlation coefficient $r = 0.951$, $P <$

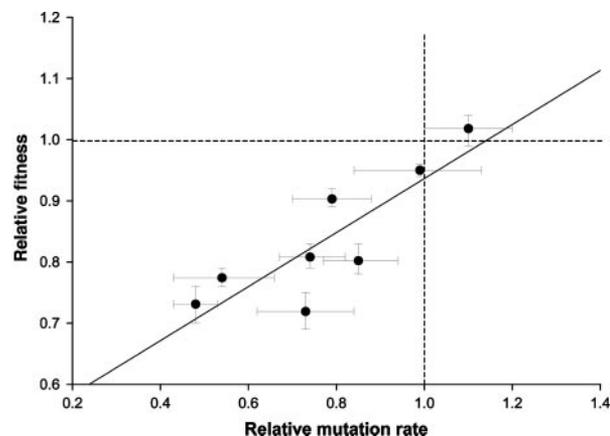


Fig. 2. Positive correlation between fitness and mutation rate, both relative to the wild type. Mean values, standard error bars, and the least-squares regression are shown.

0.0001) justifies the use of μ_{b-ML} for this latter purpose. Second, fitness relative to the wild type, measured by standard fitness assays, differed strongly among genotypes (one-way ANOVA, $P < 0.0001$), with six of eight of the mutations being significantly deleterious relative to the wild type (post-hoc test, $P < 0.005$ in all six cases). Finally, the effect on fitness of each mutation positively correlated with its effect on mutation rate ($r = 0.878$, $P = 0.002$), as shown in Fig. 2. These results demonstrate that amino acid substitutions in the *L* gene inducing an increased fidelity pay a fitness cost and suggest that high mutation rates in VSV populations might be maintained through stabilizing selection. Such a mechanism would be highly efficient, because small reductions in mutation rates would impose a heavy fitness cost.

Adaptability of Viral Clones. Under the hitchhiking hypothesis, adaptive changes are expected to appear and become fixed preferentially in those clones with a higher mutation rate. Because of the great adaptability of RNA viruses, new adaptive changes could have appeared before the competition assays, thus leading to the pattern shown in Fig. 2. To test this possibility, we evolved each genotype by performing in triplicate 10 serial infection passages in BHK₂₁ cells. If adaptability was determined by replication fidelity, a positive correlation between mutation rates and adaptation rates should be observed. Although adaptation took place in each of the 24 lineages (one-sample *t* test, $P < 0.003$ in all cases), there was no correlation between evolved fitness and ancestral mutation rates ($r = 0.076$, $P = 0.858$). Thus, our previous results (Fig. 2) are not attributable to beneficial changes arising in the populations. Further, we observed a negative correlation between mutation and adaptation rates ($r = -0.747$, $P = 0.033$), as shown in Fig. 3. This result can be understood by recalling that genotypes with low initial fitness should undergo a greater fitness gain (29), a pattern supported by our data ($r = -0.807$, $P = 0.015$). This is simply because ancestral deleterious mutations are expected to revert or to be compensated by additional changes (29–31). Therefore, our results are compatible with evolutionary reversion or compensation of mutations leading to increased fidelity, but they contradict the hypothesis of higher mutation rates conferring increased adaptation rates.

The Fitness Trade-off and Its Evolutionary Relevance. The above analysis with single-nucleotide mutants indicates the existence of a selection pressure favoring lower replication fidelity (Fig. 2). This advantage should be counterbalanced by an increasing

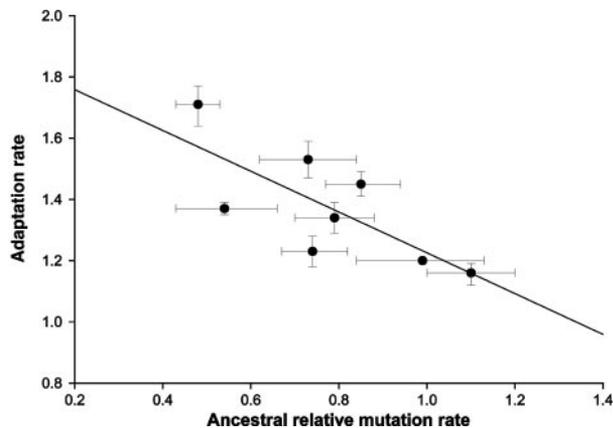


Fig. 3. Negative correlation between adaptation and ancestral mutation rates relative to the wild type. Mean values, standard error bars, and the least-squares regression are shown.

genetic load, hence producing a fitness trade-off. However, it is well conceivable that genotypes with a lowered mutation rate and unaltered fitness may exist in a small proportion. Although rare, these variants would be favored by natural selection, because they would not pay the fitness cost of replication fidelity. Consequently, adaptation could come about with no change in mutation rate. On the contrary, if the cost of replication fidelity was hardly evitable, fitness improvement should be accompanied by increased mutation rates. Hence, we decided to measure the mutation rate of the 24 evolved genotypes. We observed that the average relative mutation rate increased from 0.758 ± 0.039 in the ancestors to 1.694 ± 0.211 in the evolved genotypes (Table 2), the difference being highly significant (paired *t* test, $P < 0.001$). Genes involved in replication (*L* and *P* genes, 6,330 and 797 nucleotides long, respectively) represent 67% of the VSV coding sequence. During adaptation, changes in this region should affect both fitness and mutation rate, whereas the remaining should affect fitness exclusively. Comparing each of the 24 evolved genotypes with the ancestors (Table 2), it turned out that the mutation rate was significantly increased in 67% (16/24) of the cases. These results strongly suggest that changes in mutation rate are inevitably coupled to changes in fitness. Finally, the positive correlation between fitness and mutation rate observed in the ancestors remained significant for the evolved genotypes ($r = 0.758$, $P = 0.029$).

Table 2. Changes in mutation rate during the evolution experiment

Genotype	Relative mutation rate			
	Ancestral	Evolved		
		1	2	3
4984	1.074 ± 0.096	2.370**	0.988	1.573*
5124	0.529 ± 0.111	0.286	1.520*	1.095*
5540	0.468 ± 0.051	4.515***	0.983*	4.533***
6376	0.724 ± 0.071	2.281**	0.576	1.530**
7461	0.708 ± 0.107	1.124	1.922**	1.067
8775	0.769 ± 0.092	1.773**	1.812**	1.690*
10739	0.961 ± 0.139	2.100*	0.733	1.518
10892	0.832 ± 0.083	1.573*	0.803	2.281**
Total	0.758 ± 0.039	$1.694 \pm 0.211***$		

In all cases, a significant fitness improvement was observed. Asterisks denote statistical significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Optimal Mutation Rate. In asexuals, as a result of the above fitness trade-off, a theoretically optimal genomic mutation rate (U_{OPT}) exists, which depends solely on the function relating mutation rate to fitness, according to the following expression: $\log W = A + \gamma \log \mu$, where $\gamma = U_{OPT}$ (17). To estimate γ from our data, we need to express the mutation rate and the log-fitness in the same units. For the former, as justified above, it is preferable to use the null-class estimation (μ_0). The latter is given per hour (see *Methods*), but it can easily be expressed per replication event, because e^r individuals are produced per hour. After this transformation, the linear regression estimations are $A = 1.778 \pm 0.540$ and $\gamma = 0.160 \pm 0.047$ ($r = 0.815$, $P = 0.014$), the estimated optimal mutation rate being $U_{OPT} = 0.160$ (CI₉₅: 0.045–0.275) substitutions per genome per replication event.

Discussion

To compare estimations from different sources or to test evolutionary models, it is sometimes desirable to transform raw data into commonly used units and scales. Our mutation rate estimations are given per monoclonal antibody-resistant locus. Because six single-nucleotide substitutions conferring the resistant phenotype have been described (21), the corresponding mutation rate per nucleotidic site can be obtained as $\mu \times 3/6$ (μ being a given mutation rate estimator) and transformed to a genomic scale by multiplying per the genome length (11,162 nucleotides). Using the null-class method, this yields a genomic mutation rate of $U_0 = 0.069$ substitutions per genome per replication event for the wild-type. Taking ML values based on binary and linear replication as lower and upper bounds, respectively, we can give an interval of 0.050–0.122 mutations per genome per replication cycle. Finally, mutation rates can be converted into infection cycle units as follows: if replication was entirely linear, then an infection cycle would match a replication cycle, thus giving 0.122 mutations per genome per infection cycle. On the other hand, if replication was entirely binary, then $\log_2 K$ replication cycles would be completed in each infection cycle, with K being the per-cell viral yield. For the wild type, we used the approximation of Miralles *et al.* (25) and obtained $K = 1,250$; thus, $\log_2 K = 10.28$, and the mutation rate per genome per infection cycle would be $0.050 \times 10.28 = 0.514$. These latter estimations are interesting, because an infection cycle in viruses is the equivalent of a generation, in the sense that it is the minimum time necessary to express all phenotypic characters.

Although pioneer experiments based on limit RNase T1 cleavage (32, 33) suggested that the mutation rate in VSV could be as high as 2.8 changes per genome per replication cycle, these values were later discarded (24). Using the observed frequency of resistant genotypes against a monoclonal antibody similar to ours, Holland *et al.* (22) reported a mutant frequency of 1.7×10^{-4} . Drake and Holland (24) used these data to estimate $U = 1.52$ substitutions per genome per infection cycle, assuming an entirely linear replication. Our estimation based on this same assumption is 10-fold lower ($U = 0.122$). A possible explanation for this discrepancy might come from the fact that our wild type is a genetic chimera, in which the N-terminal region of the polymerase gene was isolated from an interfering defective particle (34). Similarly, mutation rates can strongly vary across sites and, in this sense, our estimate is probably more reliable, because it is based on six target sites, whereas the estimate made by Drake and Holland (24) was extrapolated from a single target site. In any case, all data clearly indicate that the genetic load in VSV populations (and in general, in RNA viruses) is extremely high. The mean fitness of the population (which is one minus the genetic load) can be estimated in the mutation–selection equilibrium as $\bar{W} = e^{-U}$, where U is expressed in generations (35). According to this, VSV populations undergo an 11% (using $U = 0.122$) to 78% (using $U = 1.52$) fitness reduction as a consequence of

mutation (notice that the latter U value might hence seem unrealistically high). Taking into account the quasispecies effects (36) and using Sanjuán *et al.* (19) data to estimate the selection coefficient, this burden is even stronger (46–94% using the above mutation rate values). Thus, even a moderate increase in replication fidelity would confer a substantial adaptive advantage. Our results and current evidence (see Introduction) suggest that this increase is possible from a biochemical standpoint.

Substantial work on replication fidelity has been done with DNA polymerases, generally concluding that the latter depends on the balance between exonuclease and polymerase activities (reviewed in ref. 37). However, the relationship between replication rates and fidelity has not been explored for RNA polymerases. Because they lack exonuclease activity (2), fidelity might be directly related to the rate of polymerization. This pattern is predicted by the kinetic proofreading hypothesis (38), which has already been proven to be relevant to various biological processes such as translation (39), cellular signal transduction (40), or DNA packaging (41). Under this hypothesis, error rates can be reduced by introducing an effective time delay between the formation of the enzymatic activated complex and the incorporation of the nucleotide into the nascent chain. This delay allows incorrect base pairs to dissociate before the polymerization step because of their higher off-rates, and thus correct bases are preferentially incorporated. Therefore, it

seems possible that RNA viruses could increase their replication fidelity, but this would come at a cost, because it would slow down the rate of replication. RNA viruses are characterized by enormous burst size, small genomes with frequently overlapping reading frames, fluctuating population sizes, lack of redundancy, poor metabolic pathways, and short generation times. Additionally, in VSV, as well as in other RNA viruses, replication genes deserve a considerable portion of the whole genome. In sum, RNA viruses represent an extreme form of r -selected populations (42), in which fast replication should be strongly favored and probably maintained, to the detriment of proofreading mechanisms. An optimal mutation rate is reached when this selection pressure favoring low-fidelity replication is counterbalanced by selection, favoring a lower genetic load (16, 17). According to our results, this optimal value is compatible with observed mutation rates. Therefore, without discarding the role of other evolutionary factors, high mutation rates in RNA viruses might be explained in terms of a fitness trade-off between replication fidelity and efficiency.

We thank P. J. Gerrish (Mexican Institute of Petroleum, Mexico City) for providing the software for ML mutation rate estimation as well as S. F. Elena and E. C. Holmes for critical reading of the manuscript. This work was supported by a predoctoral fellowship from the Spanish Ministerio de Educación e Ciencia (to V.F. and R.S.) and by grants from the Ministerio de Ciencia y Tecnología (Spain) and the Wellcome Trust (U.K.) (to A.M.).

- Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. (1998) *Genetics* **148**, 1667–1686.
- Steinhauer, D. A., Domingo, E. & Holland, J. J. (1992) *Gene* **122**, 281–288.
- Bakhanashvili, M., Avidan, O. & Hizi, A. (1996) *FEBS Lett.* **391**, 257–262.
- Lewis, D. A., Bebenek, K., Beard, W. A., Wilson, S. H. & Kunkel, T. A. (1999) *J. Biol. Chem.* **274**, 32924–32930.
- Wisniewski, M., Palaniappan, C., Fu, Z., Le Grice, S. F., Fay, P. & Bambara, R. A. (1999) *J. Biol. Chem.* **274**, 28175–28184.
- Boyer, P. L. & Hughes, S. H. (2000) *J. Virol.* **74**, 6494–6500.
- Pfeiffer, J. K. & Kirkegaard, K. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 7289–7294.
- Pugachev, K. V., Guirakhoo, F., Ocran, S. W., Mitchell, F., Parsons, M., Penal, C., Girakhoo, S., Pougatcheva, S. O., Arroyo, J., Trent, D. W., *et al.* (2004) *J. Virol.* **78**, 1032–1038.
- Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982) *Science* **215**, 1577–1585.
- Novella, I. S., Duarte, E. A., Elena, S. F., Moya, A., Domingo, E. & Holland, J. J. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5841–5844.
- Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. (1997) *Nature* **387**, 703–705.
- Smith, J. M. & Haigh, J. (1974) *Genet. Res.* **23**, 23–35.
- Orr, H. A. (2000) *Genetics* **155**, 961–968.
- Johnson, T. & Barton, N. H. (2002) *Genetics* **162**, 395–411.
- Kimura, M. (1967) *Genet. Res.* **9**, 23–34.
- Dawson, K. J. (1998) *J. Theor. Biol.* **194**, 143–157.
- Dawson, K. J. (1999) *Theor. Popul. Biol.* **55**, 1–22.
- Chare, E. R., Gould, E. A. & Holmes, E. C. (2003) *J. Gen. Virol.* **84**, 2691–2703.
- Sanjuán, R., Moya, A. & Elena, S. F. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 8396–8401.
- Luria, S. E. & Delbrück, M. (1943) *Genetics* **28**, 491–511.
- Holland, J. J., de la Torre, J. C., Clarke, D. K. & Duarte, E. (1991) *J. Virol.* **65**, 2960–2967.
- Holland, J. J., de la Torre, J. C., Steinhauer, D. A., Clarke, D., Duarte, E. & Domingo, E. (1989) *J. Virol.* **63**, 5030–5036.
- Stech, J., Xiong, X., Scholtissek, C. & Webster, R. G. (1999) *J. Virol.* **73**, 1878–1884.
- Drake, J. W. & Holland, J. J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 13910–13913.
- Miralles, R., Moya, A. & Elena, S. F. (2000) *J. Virol.* **74**, 3566–3571.
- Ríos, S. (1977) in *Métodos Estadísticos* (Ediciones del Castillo, Madrid), 2nd Ed., pp. 328–331.
- Wagner, R. R. & Rose, J. K. (1996) in *Virology*, eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott, Williams & Wilkins, Philadelphia), 2nd Ed., Vol. 1, pp. 1121–1137.
- Chao, L., Rang, C. U. & Wong, L. E. (2002) *J. Virol.* **76**, 3276–3281.
- Elena, S. F., Dávila, M., Novella, I. S., Holland, J. J., Domingo, E. & Moya, A. (1998) *Evolution (Lawrence, Kans.)* **52**, 309–314.
- Moore, F. B., Rozen, D. E. & Lenski, R. E. (2000) *Proc. R. Soc. London Ser. B* **267**, 515–522.
- Sanjuán, R., Cuevas, J. M., Moya, A. & Elena, S. F. (May 6, 2005) *Genetics*, 10.1534/genetics.105.040741.
- Steinhauer, D. A. & Holland, J. J. (1986) *J. Virol.* **57**, 219–228.
- Drake, J. W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4171–4175.
- Whelan, S. P., Ball, L. A., Barr, J. N. & Wertz, G. T. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 8388–8392.
- Kimura, M. & Maruyama, T. (1966) *Genetics* **54**, 1337–1351.
- Krakauer, D. C. & Plotkin, J. B. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 1405–1409.
- Goodman, M. F. & Fygenon, K. D. (1998) *Genetics* **148**, 1475–1482.
- Hopfield, J. J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4135–4139.
- Ruusala, T., Ehrenberg, M. & Kurland, C. G. (1982) *EMBO J.* **1**, 741–745.
- McKeithan, T. W. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5042–5046.
- Yan, J., Magnasco, M. O. & Marko, J. F. (1999) *Nature* **401**, 932–935.
- Pianka, E. R. (1970) *Am. Nat.* **104**, 592–597.