Resistance of virus to extinction on bottleneck passages: Study of a decaying and fluctuating pattern of fitness loss

Ester Lázaro*, Cristina Escarmís†, Juan Pérez-Mercader*‡, Susanna C. Manrubia*, and Esteban Domingo*†

*Centro de Astrobiología (Consejo Superior de Investigaciones Científicas–Instituto Nacional de Técnica Aeroespacial), Associated with the National Aeronautics and Space Administration Astrobiology Institute, Carretera de Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain; and †Centro de Biología Molecular ''Severo Ochoa'' (Consejo Superior de Investigaciones Científicas–Universidad Autónoma de Madrid), Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

Communicated by Murray Gell-Mann, Santa Fe Institute, Santa Fe, NM, May 5, 2003 (received for review February 19, 2003)

RNA viruses display high mutation rates and their populations replicate as dynamic and complex mutant distributions, termed viral quasispecies. Repeated genetic bottlenecks, which experimentally are carried out through serial plaque-to-plaque transfers of the virus, lead to fitness decrease (measured here as diminished capacity to produce infectious progeny). Here we report an analysis of fitness evolution of several low fitness foot-and-mouth disease virus clones subjected to 50 plaque-to-plaque transfers. Unexpectedly, fitness decrease, rather than being continuous and monotonic, displayed a fluctuating pattern, which was influenced by both the virus and the state of the host cell as shown by effects of recent cell passage history. The amplitude of the fluctuations increased as fitness decreased, resulting in a remarkable resistance of virus to extinction. Whereas the frequency distribution of fitness in control (independent) experiments follows a log-normal distribution, the probability of fitness values in the evolving bottlenecked populations fitted a Weibull distribution. We suggest that multiple functions of viral genomic RNA and its encoded proteins, subjected to high mutational pressure, interact with cellular components to produce this nontrivial, fluctuating pattern.

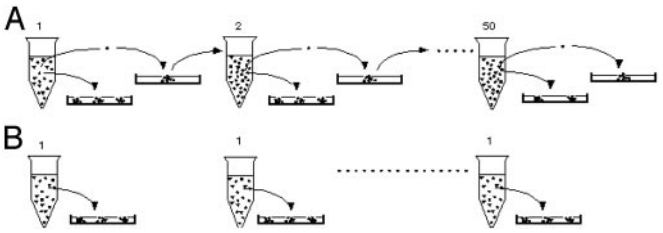

Fig. 1. Schematic representation of virus platings as carried out in this article. (*A*) Serial transfers. Clones C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10} (described in ref. 19 and in *Materials and Methods*) were diluted and plated for isolation of virus from an individual plaque (upper plates) and for titration of infectious particles (lower plates). Each viral population was derived from a plaque of the previous plating, and the serial plaque-to-plaque transfers were repeated a total of 50 times. (*B*) Control platings. C-S8c1, C_{22}^{9} p50, MARLS, and RGG were repeatedly plated to determine the influence of the cells on virus titer. Each titration was a dead end for these control platings of nonevolving virus. A total of 50 platings were carried out with aliquots of the same population (indicated as 1). Procedures for virus plating and plaque development are detailed in *Materials and Methods*.

RNA viruses mutate at rates in the range of 10^{-3} to 10^{-5} base substitutions per nucleotide copied (1, 2). These values are several orders of magnitude larger than those normally encountered during replication of viral DNA and many orders of magnitude greater than that of cellular DNA (3, 4). One of the consequences of these high mutation rates is that RNA virus populations are composed of ensembles of closely related, nonidentical genomes that are known as viral quasispecies (5–10). Viral quasispecies evolve as a result of competition, selection, and random sampling events acting on continuously arising mutant genomes. A viral quasispecies is generally dominated by one or several most-fit genomes surrounded by a mutant spectrum whose components rank according to their relative fitness in the environment in which replication takes place (7, 9, 10). Fitness is defined here as the relative replication capacity of viruses measured in growth competition experiments (11, 12).

Alterations in population size have a decisive effect in the evolution of fitness of viral quasispecies. Large population passages of RNA viruses often result in fitness gains (13–16). In contrast to this finding, repeated bottleneck passages (experimentally carried out by plaque-to-plaque transfers of a virus; compare Fig. 1) are known to result in average fitness losses of a number of different RNA viruses (17–20). Such losses have been interpreted as an accentuation of Muller's ratchet (21), or accumulation of deleterious mutations predicted to occur in asexual populations of organisms when no compensatory mechanisms such as sex or recombination operate (17–25). In the case of the important viral pathogen foot-and-mouth disease virus (FMDV), a representative of the RNA family of virus termed Picornaviridae (26), fitness losses as a result of plaque-to-plaque transfers have been associated with an accumulation of mutations at a rate of 0.3 nucleotide substitutions per transfer per genome. The process of fitness decrease, instead of being monotonic, was observed to follow a fluctuating pattern.

In this article, we report on the analysis of the statistical properties of the fluctuations in virus titer obtained with four FMDV clones subjected to serial transfers (27). To study the role played by the host cells in the process, the values of virus yield obtained in 50 independent platings were also analyzed. The results obtained show that the magnitude of the fluctuations depends on both the virus and the host cell. After the initial decrease in fitness, the system settles into a statistically stationary state (27, 28). In the case of clones subjected to serial transfers, the observed pattern of fluctuations can be described by a Weibull distribution (29), suggesting that the process belongs to a broad category of phenomena previously documented in many different fields, from physics to biology to economics (30, 31), and related to emergence (32). Our results imply that the complexity of mutational events in the virus interacting with the host cell results in considerable resistance of the virus to extinction, and have implications for the design of new antiviral strategies.

Materials and Methods

Virus, Fitness Values, Cells, and Infections. The origin and characteristics of the FMDV clones and populations used for this work have been described. FMDV C-S8c1 (33) is considered as the

Abbreviation: FMDV, foot-and-mouth disease virus.

‡To whom correspondence should be addressed. E-mail: mercader@laeff.esa.es.

www.pnas.org/cgi/doi/10.1073/pnas.1332668100

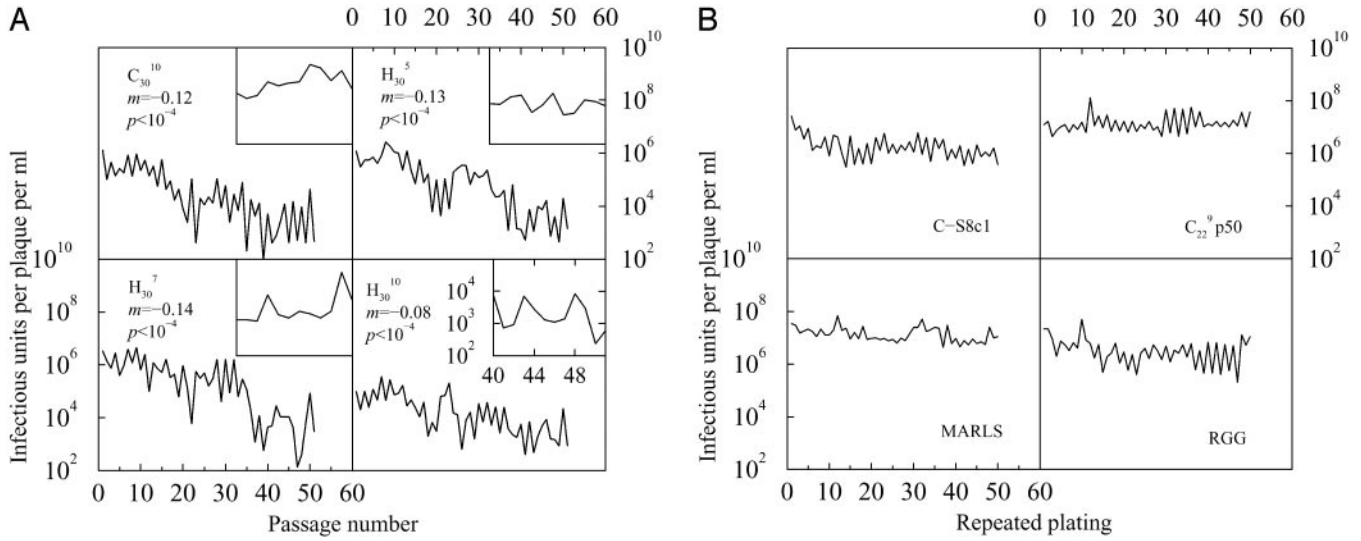

Fig. 2. Infectious units produced by sequential or repetitive platings of FMDV clones and populations. (*A*) Logarithm of the number of infectious virus progeny per ml per plaque as a function of passage number of clones C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10}. The inset in each section corresponds to the infectious units per ml produced per plaque (plaque transfers 40–50) by using cells with identical passage history (serial 1:6 dilutions). The origin of the clones is described in ref. 19 and in *Materials and Methods*, and the experimental procedure is schematically depicted in Fig. 1*A*. All main plots share the scale in both axes. The scale for insets is also shared, as represented in *Lower Right*. (*B*) Control (nonevolving) populations C-S8c1, C_{22}^{9} p50, MARLS, and RGG were repeatedly plated as schematically depicted in Fig. 1*B*. The procedures used are described in *Materials and Methods* and references therein. All four sections have the same scale as in *A*.

reference clone, with a relative fitness of 1 in BHK-21 (baby hamster kidney) cells. RGG (Arg-Gly-Gly) (34), MARLS (Mar mutant Leucine to Serine) (35, 36), and C_{22}^{9} p50 (16) are likewise used in the repeated platings of the control experiments (Fig. 1*B*), and have higher fitness than C-S8c1 (MARLS > RGG > C_{22}^{9} p50 > C-S8c1). These control populations were stored frozen in separate aliquots used for each plating. Clones C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10} (19) used in the serial transfer experiments (Fig. 1*A*) have fitness ranging from 0.1 to 0.5, which is lower than that of C-S8c1. These last four clones were obtained after subjecting C-S8c1 p2 or CS8c1 p113 to 30 plaque-to-plaque transfers (19) previous to the 50 additional passages described here.

BHK-21 cells were grown in DMEM containing 5% FCS. Cells were passaged for a maximum of 28 serial passages, and then a fresh aliquot of cells was thawed and used. For plating of virus, cells were seeded at a 1:2.5 dilution of a trypsin-treated confluent monolayer 24 h before plating and following standard procedures (refs. 16, 19, and 37 and references therein). All platings were carried out with just-confluent cell monolayers. Plaque development was for 24 h using DMEM/2% FCS/1% DEAE dextran/0.5% agar (DIFCO; ref. 19). Plating of each virus (Fig. 1) was carried out in triplicate, either in the conditions of the serial transfer experiments (Fig. 1*A*), or in the conditions of the control platings (Fig. 1*B*). The average number of progeny plaques per ml (SD never exceeded 20%) was taken as a measure of relative fitness.

Statistical Procedures. Statistical calculations were done by using the packages DATA MANIPULATION, CONTINUOUS DISTRIBUTIONS, DESCRIPTIVE STATISTICS, LINEAR REGRESSION, and NONLINEAR FIT for MATHEMATICA 3 (Wolfram Research). The values for the Kolmogorov–Smirnov statistics were determined by using the program STATGRAPHICS 7.0. The computational program for the calculation of the Hurst exponent has been written by us, and the method is described in *Appendix A*.

Results

Fitness Loss of FMDV Clones. Clones C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10}, which were derived by 30 plaque-to-plaque transfers of C-S8c1 p2 or C-S8c1 p113 (see *Materials and Methods* and ref. 19) were subjected to 50 additional plaque transfers on monolayers of BHK-21 cells. The amount of virus progeny in each cloning event was quantitated as described in *Materials and Methods* and depicted in Fig. 1*A*. Control platings involved repeated plating of the nonbottlenecked viral populations C-S8c1, C_{22}^{9} p50, MARLS, and RGG on BHK-21 cell monolayers (Fig. 1*B*). The natural logarithm of yield of progeny virus plotted as a function of passage number shows a decreasing trend that was not observed in control platings (Fig. 2). This trend (slope was estimated by regression analysis) is highly significant (*P* values are included in Fig. 2). Fluctuations in virus titer were observed in the four clones subjected to plaque-to-plaque transfers. Such fluctuations were also present, although they were less pronounced, in control platings, particularly those of the high-fitness MARLS (Fig. 2). We interpreted this as pointing out a possible role of the host cell in the viral yield. Because alternating passages of cells differed in the cell dilution at the penultimate cell passage before virus plating (1:6 versus 1:16) we explored the possibility that this difference would be the cause of the observed fluctuations. However, the time series corresponding to the number of infectious units at alternating passages also showed a similar fluctuating pattern, indicating that the cause of the fluctuations must be different. Moreover, elimination of this difference in cell passage history did not result in the disappearance of the fluctuations seen with low-fitness viruses (sections for passages 40–50 are inserted in Fig. 2*A*; see also *Discussion*). Therefore, we infer that serial plaque-to-plaque transfers of FMDV clones resulted in a decaying and fluctuating pattern on the production of infectious virus progeny.

The decreasing trend is a consequence of the sequential transfer of viral genomes, whereas fluctuations were influenced both by the virus and the host cells.

To investigate the evolution of fitness after a larger number of plaque-to-plaque passages, clone C_{30}^{10} was subjected to 50 additional passages (Fig. 3 and ref. 27). The results obtained suggest that after the initial decrease in fitness, a stationary state is reached. Results of simulations based on theoretical models (28) provide additional support for the presence of stationary states of fitness when viral populations are passed with a fixed bottleneck size (defined by the number of particles that initiate the replication at every passage).

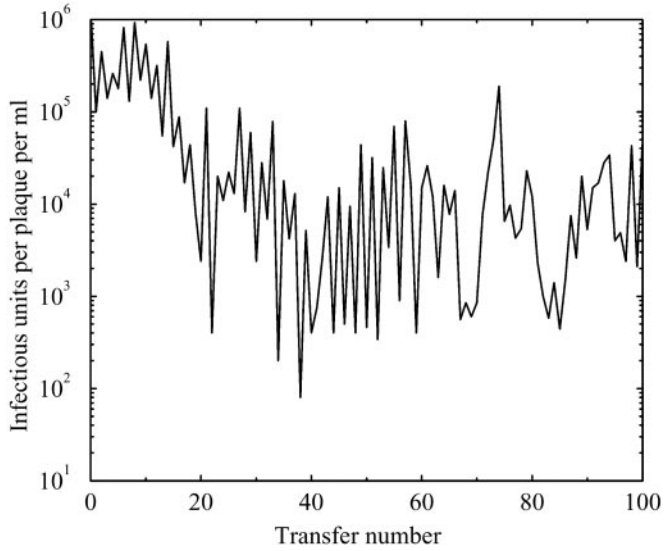

Fig. 3. Infectious units per ml per plaque produced during 100 plaque-to-plaque passages of clone C_{30}^{10}. After the initial decrease in fitness we observe that the system settles into a statistically stationary state with large fluctuations. After plaque transfer 60, the cells have an identical passage history (1:6).

Significance of the Amplitude of Fluctuations in Virus Yield. To investigate the differences between the fluctuations found in both repeated plating and serial transfers, we carried out an analysis of the statistical parameters describing experimental data.

The coefficient of variation of the logarithm of the data series (Table 1) indicate that fluctuations in the yield of infectious virus were 2.4- to 6.2-fold larger for clones C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10} subjected to serial plaque transfers than for the nonevolving, control C-S8c1, C_{22}^{9} p50, MARLS, and RGG populations. A similar conclusion was obtained for the ratios of infectivity between consecutive passages. The fluctuations in viral yield now ranged from 1.7- to 8.5-fold (Fig. 4 and Table 1).

Statistical Distributions. Often, the presence of fluctuations in the evolution of a physical system indicates that one is dealing with some form of self-organization, or with regimes where many components in the system are simultaneously involved. This finding is the case in the turbulent regime of fluid motion, where velocity fluctuations indicate that eddies of many different sizes are acting to generate the turbulent motion. In other physical systems, such as for example in Brownian motion, the presence of complicated

Table 1. Determination of the coefficient of variation of the data series

	Coefficient of variation	
Viral clone	A	B
C_{30}^{10}	0.25	0.34
H_{30}^{5}	0.22	0.22
H_{30}^{7}	0.23	0.21
H_{30}^{10}	0.19	0.21
C-S8c1	0.07	0.10
C_{22}^{9} p50	0.05	0.08
MARLS	0.04	0.04
RGG	0.08	0.12

The coefficient of variation was calculated as the ratio between the SD and the mean of the data series. A, the series corresponding to the logarithm of infectious units per ml recovered in each plaque. B, the series obtained after calculating the ratio between consecutive passages of the A series.

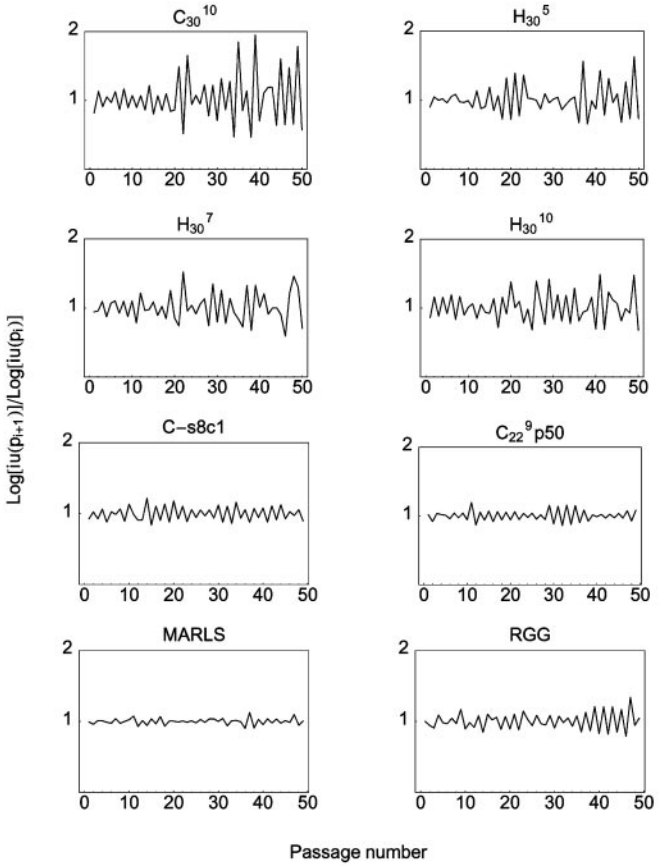

Fig. 4. Amplitude of the fluctuations of the number of infectious progeny quantified through the logarithm of infectious units per ml per viral plaque at each passage (iu at p_{i+1}) relative to the logarithm of infectious units per ml per viral plaque at the previous passage (iu at p_i). The experimental design is depicted in Fig. 1, and the origin of viruses and methods for plating on cell monolayers are detailed in *Materials and Methods*.

fluctuations is an indicator of a very deep simplicity: a Gaussian probabilistic process that describes the diffusion of small-size particles.

Inspired by the above, we have studied the fluctuations in infectious virus progeny production and the presence of long-range correlations in the process. We have applied the *rescaled range analysis technique* to our time series representing the experimental data. This method was originally introduced by H. E. Hurst (38) to analyze natural phenomena characterized by the presence of increasingly large fluctuations (compared with the typical deviations) as the length of the time series increased. It can be used to determine the nature and properties of long-time correlations. Our analysis shows that the typical values of H for all of the clones analyzed, as well as for the control series, are $\approx$0.75. Hence, we face a system with large, long-range correlated fluctuations (see *Appendix A*).

From the experimental time series for the number of infectious virus progeny per plaque as a function of passage number for each of the clones (Fig. 2), we have calculated the cumulative frequency distribution. The cumulative frequency distribution may be straightforwardly interpreted as a probability distribution function $P(Y \leq Y_0)$ that the number of infectious units per plaque, Y, is less than or equal to Y_0. Furthermore, once the probability distribution function is identified, one can then proceed to model the process in statistical terms or, alternatively, think of the underlying process as a many-body physical problem and search for an effective dynamic governing it. Here we will only attempt to identify the probability

Table 2. Fit of cumulative frequencies of the number of infectious units per ml recovered in each plaque (denoted by *Y*) to a Weibull distribution

Viral clone	α	k	Degree of significance
C_{30}^{10}	0.40 ± 0.02	5.9×10^4	0.94
H_{30}^{5}	0.55 ± 0.08	2.0×10^5	0.80
H_{30}^{7}	0.52 ± 0.05	4.2×10^5	0.61
H_{30}^{10}	0.55 ± 0.03	2.4×10^4	0.88
C-S8c1	0.99 ± 0.09	2.4×10^6	0.52
C_{22}^{9} p50	0.96 ± 0.30	1.5×10^7	0.03
MARLS	1.56 ± 0.29	1.7×10^7	0.35
RGG	0.89 ± 0.19	4.4×10^6	0.47

In this table, α and k are the two parameters characterizing the Weibull distribution function, given by $P(Y \leq Y_0) = 1 - \exp[-(Y_0/k)^{\alpha}]$. We fit the cumulative number of infectious units per ml per plaque for each of the viruses enumerated to the above expression by adjusting α and k. The goodness of the fit was evaluated through the degree of significance (P value) of the Kolmogorov–Smirnov statistic for each case.

distribution and will not present any results on a possible viral dynamics leading to it.

Once the presence of "long tails" in the data had been detected, we first considered various functional forms for the probability distribution. We tried simple power laws, as in the Pareto distribution, and then more complex probability distribution functions, such as Gaussian, log-normal, Weibull, and others. The results of fitting the experimental data to these distributions are summarized in Tables 2 and 3, where we have reported only the most significant results. We concluded that Pareto and other scale-free distributions did not fit the data because the discrepancies between data and distribution were >20%. The infectivity data clearly fell into two distinct classes: (*i*) C_{30}^{10}, H_{30}^{5}, H_{30}^{7}, and H_{30}^{10}, which had been subjected to serial plaque-to-plaque transfers, followed a pattern which is best described by a Weibull distribution; on the other hand, (*ii*) the data for the nonevolving, control populations C-S8c1, C_{22}^{9} p50, MARLS, and RGG were best described by a log-normal distribution (see Tables 2 and 3 and Fig. 5). These conclusions are also supported by Monte Carlo simulations and by the plotting of the data in rank-ordering graphics (Fig. 5), in which the Weibull distributions are well represented by straight lines (30) according to the expression

$$[Y(n)]^{\alpha} = -a \log n + b. \quad [1]$$

Table 3. Fit of cumulative frequencies of the number of infectious units per ml recovered in each plaque (denoted by *Y*) to a log-normal distribution

Viral clone	μ	σ	Degree of significance
C_{30}^{10}	9.9	2.5	0.83
H_{30}^{5}	10.9	2.4	0.50
H_{30}^{7}	11.6	2.6	0.20
H_{30}^{10}	9.4	1.8	0.91
C-S8c1	14.3	1.0	0.96
C_{22}^{9} p50	16.4	0.7	0.21
MARLS	16.4	0.6	0.35
RGG	4.9	1.2	0.60

μ and σ (mean and SD of the logarithm of infectious units per ml per plaque series) are the two parameters characterizing the log-normal cumulative distribution function given by $P(Y \leq Y_0) = 1/2[1 + \mathrm{Erf}\{(-\mu + \log Y_0)/\sqrt{2}\,\sigma\}]$. The goodness of fit was evaluated through the degree of significance (P value) of the Kolmogorov–Smirnov statistic for each case.

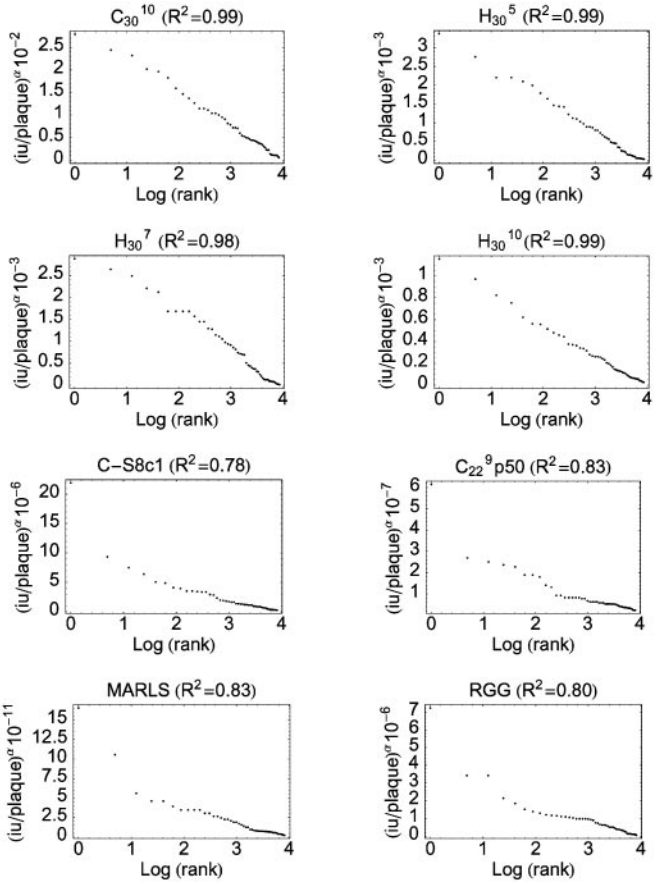

Fig. 5. Representation of the values of infectious units per ml per plaque (iu/plaque) as a function of the logarithm of the rank. The data series were ordered by decreasing values and raised to the power α that was obtained from the fitting of the frequency values to a Weibull distribution (ref. 30, see Table 2).

Here $Y(n)$ represents the number of infectious units per plaque ordered by decreasing values. The exponent α is obtained by fitting a Weibull distribution to the data; n is the rank (number of order for each value).

Because of the statistical significance of our results, it is safe to conclude that neither the evolving viral populations nor the nonevolving populations follow a logistic population dynamic, as one would expect. The dynamics of our process involve intermittent and discontinuous phenomena. For example, computational models of population dynamics, as well as empirical observations, suggest the presence of metastable states of high fitness where the system remains locked for a given time. This time shortens exponentially with increasing fitness. The dynamics we observe are therefore more complex than that described by a pure, scale-free power law. In particular, the specific nature of the phenomenon depends on the value of the exponent α and on a characteristic scale k (see Table 2).

Discussion

Unveiling the molecular mechanisms behind the extinction of viruses is of great current interest for two related reasons: (*i*) because they provide new information on the tolerance of viral genomes to accept genetic lesions while remaining replication-competent, and (*ii*) because viral extinction through increased (lethal) mutagenesis is actively investigated as a possible new antiviral strategy (10, 39–46). One of the strategies to force viruses to extinction through inefficient replication is by subjecting their quasispecies genome distributions to serial plaque-to-plaque trans-

fers that represent a severe form of repeated bottleneck, because viral population size is systematically reduced to one. As an example, several FMDV clones were brought close to extinction, although all of them were rescued by allowing longer time for plaque development or by replicating them in a liquid medium (27).

We have detected in our experiments that fitness decrease through bottleneck events, rather than being gradual and continuous, follows a fluctuating pattern influenced by both the virus and the host cell, and with an amplitude increasing as virus fitness decreases (Figs. 2 and 4). Eventually, a state where fitness fluctuations attain a statistically stationary distribution is reached (Fig. 3). Although the contribution of the host cell to the fluctuations in viral yield occurred also in control viruses repeatedly plated, the statistical properties of evolving and nonevolving viruses showed clear differences. Whereas the dynamics of the control series follows log-normal distributions, fluctuations in evolving viruses distribute according to a Weibull distribution (Table 2 and Fig. 5). We believe that the log-normal distribution has its origin in a complex dependence of the replication of viruses on the cellular state, where several steps need to be independently and successfully completed for replication to take place. Consequently, small differences are multiplicatively amplified during the course of replication and result in large fluctuations in the number of viable particles. Further studies of the underlying dynamics controlling the effect of the cellular state on replication are warranted. Because the effects of the cell were far more dramatic for low-fitness viruses (compare pattern variation in Fig. 2*A* with C-S8c1 and MARLS in Fig. 2*B*), this problem is of direct relevance to virus replication efficiency and survival, and it may relate to recent observations on the influence of cell-cycle phase in cap-dependent versus cap-independent internal ribosome entry site (IRES)-directed translation (47).

In the evolving populations, repeated bottlenecks allow for large variations in the initial state of the founder particle, an effect which is of course absent in the control series. These variations add to the cell-particle interaction and generate larger fluctuations and a "fat tail" in the probability distribution, which eventually develops a stretched exponential (Weibull) shape. The Weibull distribution has been used to describe a number of phenomena and processes in engineering and medicine (30), including cardiac contractions (48), time between events of the disease paroxysmal atrial fibrillations (49), and in modeling some epidemics (50). In our case, we believe that the simplest interpretation of our results is as a type of relaxation phenomena in which the system fluctuates out of equilibrium in response to environmental pressure: the viruses employ spontaneous mutations in a strategy to reach a more stable state. Sequence analysis of viral genomes at several intermediate passages demonstrated the accumulation of mutations in the viral genome (27), showing that the observed changes in the infectious progeny result from a synergy between genomic changes and differences in the state of the host cell.

There are strong analogies between mutation in a viral population and the notion of hopping in dispersive transport in a random environment. We can view the process of viral mutation as an evolutionary process in which the viral genome evolves from one sequence to another and where there is a replacement of dominant genomes. Because deleterious mutations are far more probable than beneficial ones, this diffusive process corresponds formally to a random walk with drift: the probability that the random walk is at position f decreases exponentially with f (51). This mechanism, together with the multiplicative replication process (which we have seen to follow a log-normal distribution function), can be used to explain the presence of fluctuations following a Weibull distribution. In fact, one can introduce a function to describe the probability of a mutation affecting fitness in a given time interval ("mutation function") and use it to study the evolution of viral populations and build models (see *Appendix B*).

Viruses are entirely dependent on host cells for their replication, in that numerous cellular proteins [surface proteins that may act as cellular receptors, protein synthesis (translation) factors, etc.], cellular RNAs and macromolecular structures and aggregates (ribosomes, membranes, etc.) participate in virus uptake, viral genome replication, and expression, as well as assembly and release of viral particles (see overviews in refs. 10 and 52). As Montroll and Shlesinger showed long ago, ". . . when a population is engaged in tasks whose completion requires the successful conclusion of many independent subtasks, the distribution function for successes in the primary task is log-normal" (53). Our results with the control viral clones completely agree with their ideas. In addition, the viral genome is a compact repository of information (10, 54): a given RNA region and also viral proteins (including unprocessed precursors) may be involved in multiple functions in such a way that variations in genetic information may trigger a cascade of perturbations sensed in the form of changes in virus yield. Thus, complex fluctuations could have their origin in the combined effect of the highly compact genetic information of the virus, in strict dependence of the host cell, and the perturbing influences of mutations in the multiple virus-cell interactions needed for completion of the virus life cycle. As the virus becomes more debilitated by deleterious mutations, compensatory mutations play a more relevant role to increase the fitness of the virus and contribute to the appearance of the fluctuating pattern.

Biologically, the large fluctuations in virus yield when viral fitness has reached very-low-fitness values (passages 40–50 in Fig. 2) underlies an extreme resistance of virus to extinction as a result of population bottlenecks and accumulation of mutations. This finding reinforces the need to search for alternative antiviral strategies not based on inhibition of replication but based instead on virus entry into error catastrophe through highly increased mutagenesis (10, 39–46). Such research using a variety of viruses warrants further study.

Appendix A: Rescaled Range Analysis

Given a time series $\{Y(t)\}$, we calculate the running sum during a time τ

$$S(\tau) = \sum_{t=0}^{\tau} Y(t), \qquad [2]$$

the average value during this interval, $\langle Y\rangle_\tau = S(\tau)/\tau$ and the associated SD $\sigma(\tau)$. The accumulated departure during this time is

$$D(t, \tau) = \sum_{i=1}^{t} \{Y(i) - \langle Y\rangle_\tau\}. \qquad [3]$$

The difference $R(\tau)$ between the maximal D_{M} and minimal D_{m} values of $D(t, \tau)$, which take place at different times in the interval $(1, \tau)$ is the expected range of variation of the accumulated departure. Empirically, Hurst (38) found out that for many phenomena the rescaled range $R(\tau)/\sigma(\tau)$ had a power-law dependence on the length τ of the observation window,

$$\frac{R(\tau)}{\sigma(\tau)} \approx \tau^H. \qquad [4]$$

The exponent $H \in [0, 1]$ is the Hurst exponent. For a simple system such as a Random Walk, $H = 0.5$, and $1/f$ noise translates into $H = 1$. Any value between 0.5 and 1 corresponds to a system with long-ranged correlations and nontrivial fluctuations, whereas values between 0 and 0.5 stand for anticorrelated time series.

Appendix B: The Mutation Function

Mutations affecting fitness, which are the ones empirically measured in this article, represent a dynamical change of state. We can define a function for the occurrence of nonneutral and nonlethal mutations, $h(t)$, as the conditional instantaneous mutation rate, so

that $h(t)dt$ is the conditional probability that the system will mutate (in a way affecting fitness) in the time interval between t and $t + dt$ given that it had not mutated until time t. Then, if T is the random time at which the mutation takes place and $f_T(t)$ is its probability density, it follows that the probability distribution function $P(T \geq t)$ is given by

$$P(T \geq t) = \int_1^\infty f_T(t')dt' = 1 - P(T \leq t). \quad [5]$$

We have seen empirically that $P(T \leq t)$ above is given by a Weibull distribution (compare Table 2). Using elementary concepts from probability theory and standard notation one can easily show that

$$h(t) = f_T(t \mid T \geq t), \quad [6]$$

$$h(t) = \frac{f_T(t)}{P_T(t)} \equiv \frac{f_T(t)}{R_T(t)}, \quad [7]$$

with $P'_T(t) = f_T(t)$. [The function of a random variable $R_T(t) = 1 - P_T(t)$ can be interpreted as the "viral resistance" to mutation.] From the above it follows that

$$f_T(t) = R_T(0)h(t)\ \exp\left[-\int_0^t h(t')dt'\right] \quad [8]$$

as well as

$$R_T(t) = (1 - P_T(0))\exp\left[-\int_0^t h(t')dt'\right]. \quad [9]$$

But, because $P_T(t) = 1 - \exp[-(t/t_0)^\alpha]$, we have for the probability density function that

$$f_T(t) = \alpha \frac{1}{t_0^\alpha} t^{\alpha-1} \exp[-(t/t_0)^\alpha], \quad \text{for } t \geq 0 \quad [10]$$

and

$$f_T(t) = 0, \quad \text{for } t < 0. \quad [11]$$

Hence the "mutation function" is found to be

$$h(t) = \alpha \frac{1}{t_0}\left(\frac{t}{t_0}\right)^{\alpha-1}. \quad [12]$$

This concept of mutation function and the associated formalism can be used to design strategies for viral extinction which we will explore elsewhere.

We thank M. Dávila and G. Gómez-Mariano for expert technical assistance, F. J. Doblas-Reyes and F. Morán for useful discussions, and an anonymous referee for constructive criticism of the manuscript. Work at the Centro de Biología Molecular Severo Ochoa was supported by BMC Grant 2001-1823-C02-01, Unión Europea Grant PSS 0884, and the Fundación Ramón Areces. S.C.M. acknowledges the Ministerio de Ciencia y Tecnología for a Ramón y Cajal Contract. Work at the Centro de Astrobiolgía was supported by the European Union, the Ministerio de Ciencia y Technología, the Comunidad Autonoma de Madrid, and the Instituto Nacional de Técnica Aeroespacial.

1. Batschelet, E., Domingo, E. & Weissmann, C. (1976) *Gene* **1,** 27–32.
2. Drake, J. W. & Holland, J. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13910–13913.
3. Drake, J. W. (1969) *Nature* **221,** 1132.
4. Drake, J. W. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7160–7164.
5. Eigen, M. (1971) *Naturwissenschaften* **58,** 465–523.
6. Eigen, M. & Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Berlin).
7. Eigen, M. & Biebricher, C. K. (1988) in *RNA Genetics*, eds. Domingo, E., Ahlquist, P. & Holland, J. J. (CRC, Boca Raton, FL), Vol. 3, pp. 211–245.
8. Domingo, E., Sabo, D., Taniguchi, T. & Weissmann, C. (1978) *Cell* **13,** 735–744.
9. Domingo, E., Holland, J. J., Biebricher, C. & Eigen, M. (1995) in *Molecular Basis of Virus Evolution*, eds. Gibbs, A., Calisher, C. & García-Arenal, F. (Cambridge Univ. Press, Cambridge, U.K.), pp. 171–180.
10. Domingo, E., Biebricher, C., Holland, J. J. & Eigen, M. (2001) *Quasispecies and RNA Virus Evolution: Principles and Consequences* (Landes Bioscience, Austin, TX).
11. Holland, J. J., de la Torre, J. C., Clarke, D. K. & Duarte, E. (1991) *J. Virol.* **65,** 2960–2967.
12. Domingo, E., Escarmís, C., Menéndez-Arias, L. & Holland, J. J. (1999) in *Origin and Evolution of Viruses*, eds. Domingo, E., Webster, R. G. & Holland, J. J. (Academic, San Diego), pp. 141–161.
13. 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.
14. Novella, I. S., Hershey, C. L., Escarmís, C., Domingo, E. & Holland, J. J. (1999) *J. Mol. Biol.* **287,** 459–465.
15. Weaver, S. C., Brault, A. C., Kang, W. & Holland, J. J. (1999) *J. Virol.* **73,** 4316–4326.
16. Escarmís, C., Dávila, M. & Domingo, E. (1999) *J. Mol. Biol.* **285,** 495–505.
17. Chao, L. (1990) *Nature* **348,** 454–455.
18. Duarte, E., Clarke, D., Moya, A., Domingo, E. & Holland, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 6015–6019.
19. Escarmís, C., Dávila, M., Charpentier, N., Bracho, A., Moya, A. & Domingo, E. (1996) *J. Mol. Biol.* **264,** 255–267.
20. Yuste, E., López-Galíndez, C. & Domingo, E. (2000) *J. Virol.* **74,** 9546–9552.
21. Muller, H. J. (1964) *Mutat. Res.* **1,** 2–9.
22. Leslie, J. F. & Vrijenhoek, R. C. (1980) *Evolution (Lawrence, Kans.)* **34,** 1105–1115.
23. Bell, G. (1988) *Sex and Death in Protozoa: The History of an Obsession* (Cambridge Univ. Press, Cambridge, U.K.).
24. Andersson, D. I. & Hughes, D. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 906–907.
25. Maynard-Smith, J. (1976). *The Evolution of Sex* (Cambridge Univ. Press, Cambridge, U.K.).
26. Racaniello, V. R. (2001) in *Fields Virology*, eds. Fields, B. N., Knipe, D. M., Howley, P. M. & Griffin, D. E. (Lippincott Williams & Williams, Philadelphia), pp. 685–722.
27. Escarmís, C., Gómez-Mariano, G., Dávila, M., Lázaro, E. & Domingo, E. (2002) *J. Mol. Biol.* **315,** 647–661.
28. Lázaro, E., Escarmís, C., Domingo, E. & Manrubia, S. C. (2002) *J. Virol.* **76,** 8675–8681.
29. Weibull, W. J. (1951) *Appl. Mech.* **18,** 293–297.
30. Laherrère, J. & Sornette, D. (1998) *Eur. Phys. J. B* **2,** 525–539.
31. Sornette, D. (2000) *Critical Phenomena in Natural Sciences* (Springer, Berlin).
32. Pérez-Mercader, J. (2001) in *Astrobiology: The Quest for the Conditions of Life*, eds. G. Horneck, G. & Baumstark-Kahn, C. (Springer, Berlin), pp. 337–359.
33. Sobrino, F., Dávila, M., Ortín, J. & Domingo, E. (1983) *Virology* **128,** 310–318.
34. Ruiz-Jarabo, C. M., Sevilla, N., Dávila, M., Gómez-Mariano, G., Baranowski, E. & Domingo, E. (1999) *J. Gen. Virol.* **80,** 1899–1909.
35. Charpentier, N., Dávila, M., Domingo, E. & Escarmís, C. (1996) *Virology* **223,** 10–18.
36. Baranowski, E., Sevilla, N., Verdaguer, N., Ruiz-Jarabo, C. M., Beck, E. & Domingo, E. (1998) *J. Virol.* **72,** 6362–6372.
37. Sobrino, F., Palma, E. L., Beck, E., Dávila, M., de la Torre, J. C., Negro, P., Villanueva, N., Ortín, J. & Domingo, E. (1986) *Gene* **50,** 149–159.
38. Hurst, H. E. (1951) *Trans. Am. Soc. Civ. Eng.* **116,** 770–799.
39. Holland, J. J., Domingo, E., de la Torre, J. C. & Steinhauer, D. A. (1990) *J. Virol.* **64,** 3960–3962.
40. Holland, J. J., De La Torre, J. C. & Steinhauer, D. A. (1992) *Curr. Top. Microbiol. Immunol.* **176,** 1–20.
41. Ji, J., Hoffmann, J. S. & Loeb, L. (1994) *Nucleic Acids Res.* **22,** 47–52.
42. Loeb, L. A., Essigmann, J. M., Kazazi, F., Zhang, J., Rose, K. D. & Mullins, J. I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 1492–1497.
43. Loeb, L. A. & Mullins, J. I. (2000) *AIDS Res. Hum. Retroviruses* **13,** 1–3.
44. Sierra, S., Dávila, M., Lowenstein, P. R. & Domingo, E. (2000) *J. Virol.* **74,** 8316–8323.
45. Grande-Pérez, A., Sierra, S., Castro, M. G., Domingo, D. & Lowenstein, P. R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 12938–12943.
46. Pariente, N., Sierra, S., Lowenstein, P. R. & Domingo, E. (2001) *J. Virol.* **75,** 9723–9730.
47. Sachs, A. B. (2000) *Cell* **101,** 243–245.
48. Araki, J., Matsubara, H., Shimizu, J., Mikane, T., Mohri, S., Mizuno, J., Takaki, M., Ohe, T., Hirakawa, M. & Suga, H. (1999) *Am. J. Physiol.* **277,** H1940–H1945.
49. Rose, M. S., Gillis, A. M. & Sheldon, R. S. (1999) *Stat. Med.* **18,** 139–154.
50. Cooley, P. C., Myers, L. E. & Hamill, D. N. (1996) *Eur. J. Epidemiol.* **12,** 229–235.
51. Feller, W. (1957) *An Introduction to Probability Theory and Its Applications* (Wiley, New York).
52. Flint, S. J., Enquist, L. W., Krug, R. M., Racaniello, V. R. & Skalka, A. M. (2000) *Virology, Molecular Biology, Pathogenesis, and Control* (Am. Soc. Microbiol., Washington, DC).
53. Montroll, E. W. & Shlesinger, M. F. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 3380–3383.
54. Gromeier, M., Wimmer, E. & Gorbalenya, A. E. (1999) in *Origin and Evolution of Viruses*, eds. Domingo, E., Webster, R. G. & Holland, J. J. (Academic, San Diego), pp. 287–343.