THE ROYAL SOCIETY

Received 3 April 2002
Accepted 5 June 2002
Published online 28 August 2002

Procedures for reliable estimation of viral fitness from time-series data

Sebastian Bonhoeffer[1*], Andrew D. Barbour[2] and Rob J. De Boer[3]

[1]*Ecology and Evolution, ETH Zurich, ETH Zentrum NW, CH-8092 Zurich, Switzerland*
[2]*Institute of Mathematics, University of Zurich, CH-8057 Zurich, Switzerland*
[3]*Theoretical Biology, Utrecht University, 3584 CH Utrecht, The Netherlands*

In order to develop a better understanding of the evolutionary dynamics of HIV drug resistance, it is necessary to quantify accurately the *in vivo* fitness costs of resistance mutations. However, the reliable estimation of such fitness costs is riddled with both theoretical and experimental difficulties. Experimental fitness assays typically suffer from the shortcoming that they are based on *in vitro* data. Fitness estimates based on the mathematical analysis of *in vivo* data, however, are often questionable because the underlying assumptions are not fulfilled. In particular, the assumption that the replication rate of the virus population is constant in time is frequently grossly violated. By extending recent work of Marée and colleagues, we present here a new approach that corrects for time-dependent viral replication in time-series data for growth competition of mutants. This approach allows a reliable estimation of the relative replicative capacity (with confidence intervals) of two competing virus variants growing within the same patient, using longitudinal data for the total plasma virus load, the relative frequency of the two variants and the death rate of infected cells. We assess the accuracy of our method using computer-generated data. An implementation of the developed method is freely accessible on the Web (http://www.eco.ethz.ch/fitness.html).

Keywords: HIV-1; fitness; selection coefficient; competition experiments; parameter estimation

1. INTRODUCTION

As evolutionary biologists will readily attest that measuring the fitness of organisms in their natural environment is notoriously difficult, but the quantification of fitness is the key to the understanding of the dynamics of evolutionary adaptation. HIV, because of its high mutation rate (Mansky 1996) and its high rate of turnover (Wei *et al.* 1995; Ho *et al.* 1995), has a remarkable capacity for rapid evolutionary adaptation, as is evidenced by its ability to escape from specific immunity (Phillips *et al.* 1991; Borrow *et al.* 1997; Goulder *et al.* 1997, 2001) or to evolve resistance to retroviral inhibitors (Larder & Kemp 1989; Larder *et al.* 1989; Boucher *et al.* 1990; Richman 1990; StClair *et al.* 1991; Ho *et al.* 1994; Richman *et al.* 1994). To predict the kinetics of evolutionary adaptation and ultimately to improve the prescription of effective therapy, it is thus necessary to quantify fitness differences between virus variants. The increasing body of literature on viral fitness, on the one hand, clearly demonstrates the awareness in the field for the need of accurate quantitative fitness estimates, but on the other hand, it reveals a shortage of reliable mathematical procedures to estimate fitness and an unawareness of the assumptions underlying the methods currently used.

Many methods for quantifying viral fitness are based on the relative growth kinetics of two virus variants growing in competition either *in vivo* (Chao 1990; Goudsmit *et al.* 1996, 1997; Eastman *et al.* 1998; Zennou *et al.* 1998) or *in vitro* (Holland *et al.* 1991; Martinez *et al.* 1991; Croteau *et al.* 1997; Harrigan *et al.* 1998; Martinez-Picado *et al.* 1999, 2000; Yuste *et al.* 1999). In such studies, typically the ratio of the two variants is plotted logarithmically against time, and the resulting slope is used as a measure of fitness. As has been pointed out before (Marée *et al.* 2000), the underlying theory shows that this slope measures the absolute fitness difference and not, as is frequently assumed, the selection coefficient or the relative fitness of the two variants. A key assumption behind this procedure to calculate absolute fitness differences is that the replication rates of the virus variants are constant in time, or, equivalently, that the virus variants grow or decline exponentially. However, this assumption is often grossly violated both *in vivo* and *in vitro*, because many factors that affect the replication rate, such as the density of susceptible target cells, may change considerably over time.

Even under conditions in which the replication rate of the viral population can safely be assumed constant, the quantification of fitness in terms of absolute fitness differences is of limited use, as it does not allow a direct comparison between different experimental set-ups or patients. This can be seen as follows: the replication rates of wild-type virus, r, and mutant virus, r', relate to each other as $r' = (1 + s)r$, where s is the selection coefficient as defined in population genetics (Nagylaki 1992). The absolute fitness difference is given by $r' - r = rs$, and hence is proportional to r. The replication rate r, however, depends on factors such as the target cell density, which may vary between patients or experiments. Consequently, the absolute fitness difference is a fitness measure that depends on the specific growth conditions and is of limited use for comparison between patients or experiments.

A more useful measure of viral fitness is the selection coefficient, $s = 1 - r'/r$, which measures the replication rates of the wild-type and mutant virus relative to each other, and thus eliminates all factors that affect the repli-

*Author for correspondence (bonhoeffer@eco.umnw.ethz.ch).

cation rate of both virus variants. To calculate the selection coefficient s from the absolute fitness difference, an estimate is needed for the replication rate r under the given experimental conditions. Because this is usually not available, many researchers have simply divided the absolute fitness difference by the generation time of the virus (Goudsmit *et al.* 1996, 1997; Harrigan *et al.* 1998; Martinez-Picado *et al.* 1999, 2000) to obtain an estimate for the selection coefficient. However, this calculation makes yet another assumption that is frequently violated. Dividing the absolute fitness by the generation time T_g only yields the selection coefficient if $1/T_g = r$. However, the reciprocal of the generation time only equals the replication rate in populations at equilibrium. While the assumption of a virus population at equilibrium can sometimes be justified *in vivo* (Goudsmit *et al.* 1996, 1997), it is certainly questionable for the standard growth competition assays *in vitro* (Harrigan *et al.* 1998; Martinez-Picado *et al.* 1999, 2000). Moreover, using the estimate for the *in vivo* generation time of $T_g = 2.6$ (Perelson *et al.* 1996) for *in vitro* growth competition experiments is inappropriate, because its reciprocal value does not generally equal the replication rate r realized in the particular experimental set-up *in vitro*.

In order to overcome the difficulties in obtaining accurate estimates of the fitness effects of individual mutations, we develop here a method that corrects for time-dependent changes of the replication rate of the virus population and reliably estimates the selection coefficient, based on an extension of a recently published approach by Marée *et al.* (2000).

2. METHODS

(a) *Estimation of selection coefficient from two time-points*

As in Marée *et al.* (2000), we describe the dynamics of competition between wild-type (wt) and mutant (mt) virus by the following set of differential equations:

$$\mathrm{d}W(t)/\mathrm{d}t = [r(t) - \delta]\,W(t), \tag{2.1}$$

$$\mathrm{d}M(t)/\mathrm{d}t = [(1 + s)r(t) - \delta]M(t). \tag{2.2}$$

The variables $W(t)$ and $M(t)$ denote the densities of cells infected with wild-type and mutant virus, respectively, at time t. The time-dependent function $r(t)$, referred to here as the replication rate, describes the growth rate of the infected cell population per infected cell. The replication rate is assumed to be time dependent, because it depends on factors such as the availability of susceptible target cells that may change with time. The parameter δ represents the death rate of virus-producing infected cells. The factor $1 + s$ denotes the replicative capacity of the mutant relative to the wild-type, where the parameter s is the selection coefficient as used in population genetics (Nagylaki 1992). The parameter s has also been referred to in the virus dynamics literature as the selective advantage (if $s > 0$) or disadvantage (if $s < 0$) (Bonhoeffer *et al.* 1997*a*,*b*; Ribeiro *et al.* 1998; Ribeiro & Bonhoeffer 2000). Note that the model assumes that the difference between wt and mt virus is manifest in the replication rate rather than in the death rate. Furthermore, the model assumes that the growth rate of both populations is proportional to the corresponding infected cell densities. The rationale behind this assumption is that the dynamics of free virus are typically fast in comparison with the infected cells (Perelson *et al.* 1996; Nowak *et al.* 1996; Ramratnam *et al.* 1999) and thus free virus can be assumed to be proportional to infected cells at all times. In what follows, we therefore do not explicitly distinguish between free virus load and infected cell load. For further discussion of the underlying assumptions see § 4.

Substituting the log mt/wt ratio $h = \ln(M/W)$ and the log wt virus load $w = \ln(W)$ in equations (2.1) and (2.2), Marée *et al.* (2000) derived the following relationship between the selection coefficient and the values of h and w at the time-points $t = 0$ and $t = T$:

$$h(T) - h(0) = s(w(T) - w(0) + \delta T). \tag{2.3}$$

The key feature of this equation is that it allows an estimation of the selection coefficient s without explicit knowledge of the replication rate $r(t)$. The effect of a time-dependent replication rate enters into equation (2.3) as a time average, as it can be shown that the average rate of replication over the time interval T is given by $(w(T) - w(0) + \delta T)/T$.

The major advantage of estimating s based on equation (2.3) is that it circumvents the problem of a time- and patient-dependent replication rate, by making use of data for the growth of the wild-type virus population. However, the method also has some shortcomings for practical use. First, the estimation of the selection coefficient is based on two time-points only. Second, equation (2.3) offers no statistical information regarding confidence intervals for the estimated selection coefficient. In the following section, we discuss how these shortcomings can be overcome.

(b) *Estimation of selection coefficient from time-series*

Reformulating equation (2.3), we have the following relationship between the data of pairs of successive observations

$$h_i = h_{i-1} + s(w_i - w_{i-1} + \delta t_i), \tag{2.4}$$

where T_i is the time of the ith observation, $t_i = T_i - T_{i-1}$, $h_i = h(T_i)$, and $w_i = w(T_i)$. Iterating this equation, we obtain for the relationship between time-points i and 0

$$h_i = h_0 + s\tau_i, \tag{2.5}$$

where $\tau_i = (\delta T_i + w_i - w_0)$ and $T_i = \Sigma_{k=1}^{i} t_k$. Hereafter, we refer to τ_i as rescaled time, although in a strict sense τ_i does not represent a time, as it is dimensionless.

Given experimental data for the fraction of mutant virus $\hat{f}_i$, the total virus load $\hat{V}_i$ and the death rate of virus producing cells δ, we can calculate the expected values of the rescaled time $\hat{\tau}_i$, and the log mt/wt ratio $\hat{h}_i$ (see Appendix A). Substituting $\hat{\tau}_i$ for τ_i in equation (2.5), we can calculate h_i for a given choice of the selection coefficient s and the initial log mt/wt ratio h_0. To measure how well the data fit the model (equation (2.5)), we use the penalty function

$$\rho(s,h_0) = \sum_{i=1}^{n} \frac{(h_0 + s\hat{\tau}_i - \hat{h}_i)^2}{\mathrm{Var}(h_0 + s\hat{\tau}_i - \hat{h}_i)}. \tag{2.6}$$

This penalty function corresponds to the total sum of squares of a linear regression but weighted by the variance of $(h_0 + s\hat{\tau}_i - \hat{h}_i)$. In contrast to the usual linear regression, both coordinates (i.e. τ and h) are subject to experimental error, because the calculation of $\hat{\tau}_i$ and $\hat{h}_i$ is based on $\hat{f}_i$, $\hat{V}_i$ and δ, which themselves are all subject to experimental error. The values of s and h_0 that best fit the data are determined by minimizing the pen-

alty function. However, in contrast to a linear regression, the best fit values of s and h_0 can only be determined numerically, because the variance in the denominator of equation (2.6) is a nonlinear function of s.

The assumptions underlying the above procedure for estimating the selection coefficient are similar to those underlying linear regression. In particular, it is assumed that the residuals $(\hat{h}_i - h_i)$ and $(\hat{\tau}_i - \tau_i)$ and δ have finite variances.

(c) *Methods for estimation of* s

We compare three methods of estimating the selection coefficient from experimental data. The first method, called here the conventional method (CM), is equivalent to the methods currently used in the literature (Goudsmit *et al.* 1996, 1997; Harrigan *et al.* 1998; Martinez-Picado *et al.* 1999, 2000) in which the slope of the regression is rescaled by generation time. CM estimates the selection coefficient by regressing the log mt/wt ratio against time and dividing the resulting regression slope by the death rate of infected cells. As pointed out in § 1, the slope of the regression of log mt/wt against time measures the absolute fitness difference given by rs. If the virus population is to a good approximation in equilibrium, then the replication rate approximately equals the death rate, i.e. $r \approx \delta$. Hence, dividing the slope of the regression rs by the death rate δ we obtain an estimate for the selection coefficient. The advantage of CM is that it requires only data for the log mt/wt ratio. Therefore the selection coefficient can be estimated in the absence of data for the total virus load. The shortcoming is that the underlying assumptions of constant replication rate and constant population size are frequently not fulfilled.

The second method, called the growth-corrected method (GM), estimates the selection coefficient based on penalty function equation (2.6). Because the calculation of the τ_i requires knowledge of the log wt virus load, the estimation of the selection coefficient with GM requires data on both virus load and mutant frequency. However, the advantage over CM is that GM corrects for the potential effects of time-dependent changes in the replication rate. The disadvantage is that routines to determine the best fit of a straight line to data with errors in both coordinates are not readily available in standard statistical packages. Therefore we have made an implementation of GM freely accessible on the Web (http://www.eco.ethz.ch/fitness.html).

Finally, the third method, called the average method (AM), computes the selection coefficient between all pairs of successive time-points according to Marée *et al.* (2000) (see equation (2.3)) and reports the mean and variance of these estimates. This method serves as a comparison with the other methods, to assess the improvement of the estimation of the selection coefficient through using a regression procedure.

(d) *Simulation of data*

To evaluate the accuracy of the estimation of the selection coefficient by the three methods described above, we generated data using the following procedure. We performed numerical simulations for the replication dynamics of wild-type and mutant virus according to equations (2.1) and (2.2), with varying parameters δ and s and different choices of the time-dependent replication rate function $r(t)$. At selected time-points, T_i, the fraction of mutants $f_i = M(T_i)/(M(T_i) + W(T_i))$ and the total virus load $V_i = M(T_i) + W(T_i)$ were stored and used to generate Gauss-distributed data for $\hat{f}_i$ and $\hat{V}_i$ with the corresponding means f_i and V_i, and chosen variance $\sigma^2_{f_i}$ and $\sigma^2_{V_i}$. Of the data thus generated, we took the longest stretch of consecutive time-points for which the fraction of mutant virus $\hat{f}_i$ is between σ_{f_i} and $1 - \sigma_{f_i}$. This selection procedure avoids a bias in the estimation of the selection coefficient that may arise, because if the fraction of mutants is close to 1 or 0, the errors are not symmetrically distributed around the mean. On the basis of these data, we then determined the expected values and variances of $\hat{h}_i$ and $\hat{\tau}_i$ according to equations (A 1)–(A 4) (see Appendix A) and used these to estimate the selection coefficient using all three methods.

3. RESULTS

Figure 1 illustrates the estimation of the selection coefficient by the three methods CM, GM and AM based on simulated datasets. The replication rate function chosen for the simulation of the data decreases as the total virus load increases. This simulates a decreased availability of target cells as the virus load increases. The comparison between the selection coefficient used for the simulation of the data and those estimated by the different methods shows that CM can lead to grossly inadequate estimates of the selection coefficient.

To assess more systematically how well the three methods estimate the selection coefficient, we performed extensive computer simulations. Table 1 shows the mean estimate of the selection coefficient $\bar{s}_{\mathrm{est}}$ based on 100 simulated datasets for all methods and different choices of the parameter values and replication rate function. Both AM and GM yield values of $\bar{s}_{\mathrm{est}}$ that are not significantly different from the selection coefficient s_{real} used for the generation of the data. Also, comparing the estimated s with s_{real} on an individual basis for each of the 100 generated datasets shows that GM and AM yield estimates of s that are in most cases not significantly different from s_{real} based on the estimated standard deviation σ_s of s. CM, by contrast, almost always fails on this account. Importantly, it consistently yields poor estimates for s even when the underlying assumption, that the replication rate of the virus population is constant in time, is fulfilled (see the replication rate function 1 in table 1). The reason for the poor estimates is that the other underlying assumption, namely that death rate approximately equals the replication rate, is not justified here (see § 2). This illustrates how CM can yield flawed results, if the underlying assumptions are violated.

GM generally leads to good estimates of s. Table 1 shows that the GM estimates a selection coefficient that is significantly different from s_{real} for less than 5 out of 100 simulations. This indicates that GM slightly overestimates the standard deviation of the estimate of s.

AM rarely leads to estimates of s which are significantly different from s_{real}. However, this is not because AM yields estimates that are close to s_{real}, but because the estimates of s have large standard deviation σ_s. Occasionally the method yields estimates that are far from the s_{real}, resulting in a large standard deviation of $\bar{s}_{\mathrm{est}}$ in some places in table 1. This reveals a weakness of the method; if, by chance, $\delta t_i \approx w_{i-1} - w_i$ for successive time-points, the method may generate outliers, because the calculation of s involves division by a number close to zero (see equation (2.4)).

Overall, the data in table 1 indicate that, based on the accuracy of the estimate of s and the conservative estimate of the corresponding standard deviation σ_s, GM is

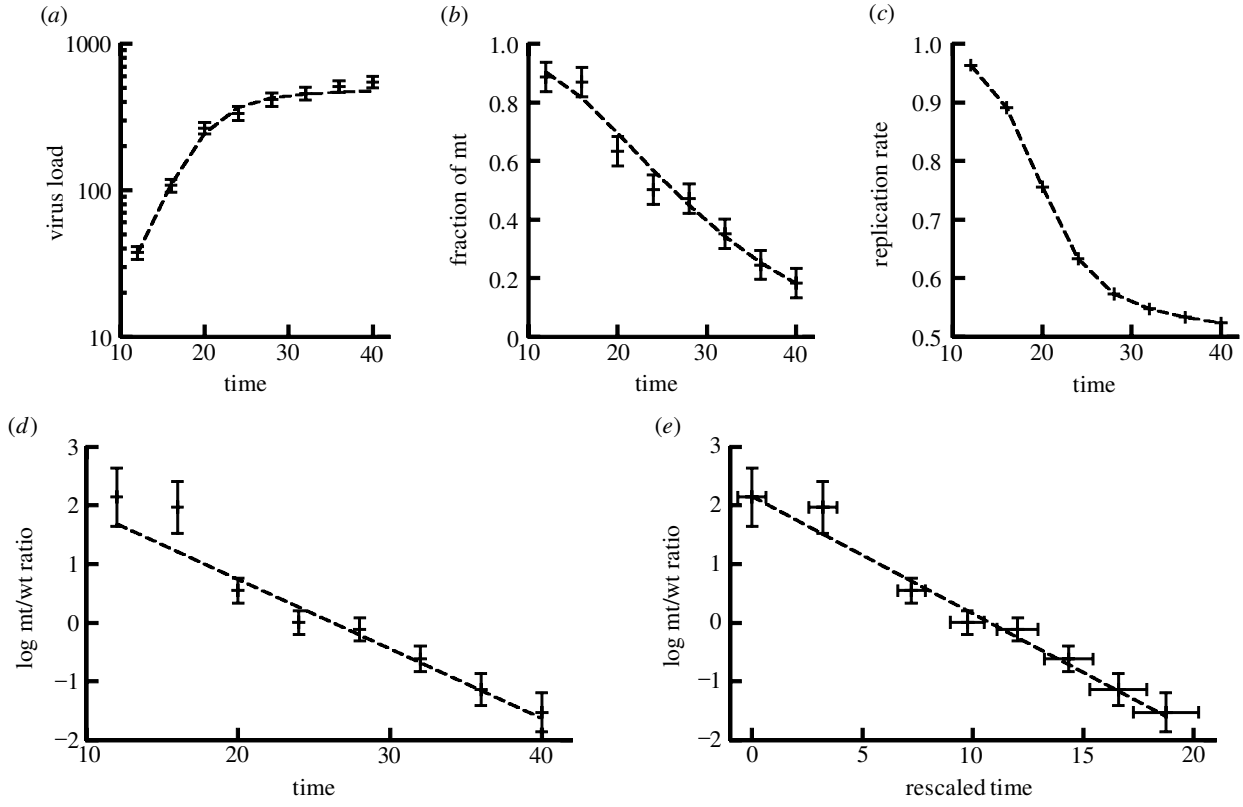

Figure 1. Comparison of the CM, GM and AM methods for the estimation of the selection coefficient based on computer-generated data. (*a*,*b*) The total virus load and the fraction of mutants generated by computer simulation as described in the main text. The dashed lines indicate the numerical solutions of equations (2.1) and (2.2), which were used to generate data with Gauss-distributed mean and chosen standard deviation (represented by points with error bars) as described in the main text. The parameters for the simulation are $s = -0.2$, $\delta = 0.5$, $\sigma_{V_i} = 0.1\,V_i$, and $\sigma_{f_i} = 0.05$. For the replication rate, we chose $r(t) = 1 - (W(t) + M(t))/1000$ (see (*c*)). The replication rate thus decreases with increasing virus load, which could for example be due to a decreased target cell availability. For the calculation of $\hat{h}_i$ and $\hat{\tau}_i$ (and their variances, see equations (A 1)–(A 4)) we used $\sigma_\delta = 0.05$. (*d*,*e*) The estimation of the selection coefficient by CM (*d*) and GM (*e*). The corresponding estimate for the selection coefficient (± s.d.) by CM is $s = -0.239 \pm 0.001$ and by GM is $s = -0.200 \pm 0.023$. For this dataset, the selection coefficient estimated by CM is significantly different ($p < 0.001$, *t*-test) from the 'real' selection coefficient $s = -0.2$ used to generate the data. Nevertheless, the goodness-of-fit probability, q, describing the probability that a fit as poor as this occurs by chance, is $q = 0.35$, indicating a deceptively good fit. The selection coefficient estimated by GM is not significantly different ($p > 0.5$) from the 'real' selection coefficient (q-value 0.79). The estimated initial log mt/wt ratio for CM is $h_0 = 3.1 \pm 0.4$ at $t = 0$ and for GM is $h_0 = 2.2 \pm 0.3$ at $\tau_0 = 0$. AM (not shown in the figure) estimates a selection coefficient of $s = -0.182 \pm 0.105$, which is also not significantly different ($p > 0.5$) from the 'real' selection coefficient. Note, however, that the standard deviation of this estimate is four- to fivefold higher than that of GM.

superior to the other methods for obtaining reliable estimates of the selection coefficient and its standard deviation. Generally, the estimates for the selection coefficient tend to be somewhat smaller (in absolute value) than s_{real} for all methods. The possible reason for this lies in the selection of the data used for the estimation procedures. As discussed in § 2d, we select from the simulated data the longest stretch of time-points for which all data for the fraction of mutants, f_i, lies between σ_f and $1 - \sigma_f$. Values close to both cut-offs may sometimes reflect time-points where the true value of f is in fact outside the range, but because of the generated random error, the simulated f lies inside the range. This may lead to a bias in the data that goes in the direction the observed systematic discrepancy between the estimated and the real selection coefficient.

4. DISCUSSION

The extensive numerical tests based on computer generated data clearly reveal the strengths and weaknesses of the different methods of estimation of the selection coefficient. CM, which estimates the selection coefficient from the slope of the logarithmic mt/wt ratio scaled by the generation time of the virus, typically yields grossly inaccurate estimates. The reason for the poor performance of this method is that the underlying assumptions of a time-constant replication rate and a total virus population in steady state are frequently violated. Given that these assumptions will frequently also be violated in real biological systems, this argues that the results of this method (and related methods used in the literature) need to be interpreted with considerable caution.

GM is based on an extension of the approach by Marée *et al.* (2000), which allows an estimation of a selection coefficient between two time-points using data for the total virus population growth to correct for a time- (and experiment-) dependent replication rate. It estimates the selection coefficient from the slope of the logarithmic ratio of mt to wt virus, whereby time is rescaled according to the total virus population growth. GM is thus applicable

Table 1. Statistical comparison of the CM, AM and GM methods (see § 2) with simulated datasets.
(Each block of three lines represents the estimation of the selection coefficient by the three methods based on 100 independent simulated datasets with, on average, eight observations for a given death rate, δ, total time T and three replication rate functions specified below. The headers of the columns are as follows: s_{real} for the 'real' selection coefficient used to generate the simulated datasets; $\bar{s}_{est}$ for the mean over 100 estimated selection coefficients; $\sigma_{s_{est}}$ for the standard error of the mean; and m for the number of times that a t-test between the estimated s and s_{real} indicated a significant difference, based on the standard deviation σ_s of the individual estimate of s. The superscripts 1, 2 and 3 indicate that differences between s_{real} and $\bar{s}_{est}$ were significant based on a t-test at a level $p < 0.01$, $p < 0.005$ and $p < 0.001$, respectively. Replication rate function 1 is given by $r(t) = 2\delta$ and is thus constant in time. Replication rate function 2 is given by $r(t) = 2\delta(1 - (W(t) + M(t))/1000)$, and describes a replication rate that decreases with increasing virus load, as would, for example, be expected under target cell limitation. Replication rate function 3 is given by $r(t) = \delta(1 + \sin(12t/T))$. This function has no particular biological motivation. Rather it was chosen to test the performance of the method under extreme fluctuations of the replication rate with time. In all simulations, we used $\sigma_\delta = 0.1\delta$, $\sigma_{f_i} = 0.05$ and $\sigma_{V_i} = 0.1 V_i$.)

		replication rate function 1			replication rate function 2			replication rate function 3		
method	s_{real}	$\bar{s}_{est}$	$\sigma_{s_{est}}$	m	$\bar{s}_{est}$	$\sigma_{s_{est}}$	m	$\bar{s}_{est}$	$\sigma_{s_{est}}$	m
parameters: $\delta = 0.5$, $T = 40$										
CM	−0.3	-0.563^2	0.083	100	-0.511^2	0.067	100	−0.286	0.039	96
GM	−0.3	−0.287	0.034	3	−0.288	0.029	5	−0.296	0.026	1
AM	−0.3	−0.254	0.056	2	−0.251	0.052	0	−0.311	1.086	0
parameters: $\delta = 0.5$, $T = 50$										
CM	−0.2	-0.381^3	0.046	100	−0.234	0.018	99	−0.176	0.018	96
GM	−0.2	−0.192	0.021	3	−0.194	0.018	6	−0.194	0.016	2
AM	−0.2	−0.17	0.028	2	−0.178	0.029	0	−0.317	0.475	0
parameters: $\delta = 0.5$, $T = 80$										
CM	−0.1	-0.191^3	0.021	100	−0.108	0.013	97	−0.094	0.011	99
GM	−0.1	−0.096	0.01	3	−0.1	0.011	0	−0.098	0.01	2
AM	−0.1	−0.091	0.016	0	−0.097	0.013	0	−0.06	0.114	0
parameters: $\delta = 0.5$, $T = 150$										
CM	−0.05	-0.093^3	0.013	100	-0.092^2	0.013	100	-0.092^2	0.013	100
GM	−0.05	−0.047	0.006	7	−0.046	0.006	6	−0.046	0.006	6
AM	−0.05	−0.042	0.007	1	−0.043	0.008	2	−0.042	0.008	2
parameters: $\delta = 1$, $T = 50$										
CM	−0.1	-0.193^3	0.024	100	−0.105	0.01	96	−0.092	0.009	100
GM	−0.1	−0.096	0.011	6	−0.098	0.009	2	−0.099	0.009	0
AM	−0.1	−0.087	0.016	2	−0.094	0.012	0	−0.253	0.81	0
parameters: $\delta = 0.1$, $T = 500$										
CM	−0.1	-0.189^3	0.021	100	−0.103	0.012	96	−0.089	0.009	96
GM	−0.1	−0.095	0.01	5	−0.096	0.011	5	−0.097	0.01	3
AM	−0.1	−0.088	0.014	0	−0.09	0.012	0	−0.074	0.088	0

to growth competition of virus variants in growing, declining or fluctuating total virus populations. The numerical tests show that GM reliably estimates the selection coefficient for a variety of choices of time-dependent replication rate functions. GM takes the variances and covariances of the log mt/wt ratio and the rescaled time into account. In contrast to CM, it cannot be performed easily with simple stastistical software packages.

AM is also based on the approach of Marée *et al.* (2000), but unlike GM it does not involve a linear-regression-type procedure. AM calculates a mean selection coefficient by averaging over the selection coefficient between all pairs of successive time-points determined by Marée *et al.*'s method. It thus serves as a comparison to the growth-correcting methods to illustrate the gain in accuracy of the estimate of the selection coefficient by a linear-regression-type procedure. However, AM can generate grossly inaccurate estimates of the selection coefficient, because statistical fluctuations in the total virus population can result in numerical problems due to division by numbers close to zero.

Overall, the numerical tests of the methods indicate that GM yields the most reliable estimates of the selection coefficient and its standard deviation. The estimated selection coefficient is typically significantly different from the selection coefficient used to generate the data in less than 5 out of 100 datasets (see table 1), indicating that this method slightly conservatively overestimates the standard deviation of the estimated selection coefficient. CM and AM often yield poor estimates. Moreover, CM consistently underestimates the standard deviation of the selection coefficient, and thus yields estimates that are highly significantly different from the selection coefficient used to generate the data (see table 1).

GM requires longitudinal data for the fraction of mutant virus and the total virus load, as well as the death rate of infected cells. Although the method was primarily developed to estimate selection coefficients from *in vivo* data, it can also be applied to *in vitro* growth competition experiments, provided the appropriate data are available. Note, however, that in this case the death rate of infected cells *in vitro* has to be known for the specific experimental set-up (see Appendix A).

The methods described in this paper are designed

specifically to measure the selective effect of drug resistance mutations in HIV. They all are based on three key assumptions underlying equations (2.1) and (2.2). The first is that the selective effect of the drug resistance mutations is manifest in differences in the intrinsic replication rate (rather than the death rate) of the virus. The rationale for this assumption is that drug resistance mutations in HIV affect the reverse transcriptase and protease and therefore affect the infectivity of the virus, rather than the death rate of infected cells. The second assumption is that the death rate can be assumed to be constant in time. This assumption is frequently made in models of virus dynamics (Ho *et al.* 1995; Nowak *et al.* 1995; Wei *et al.* 1995; Perelson *et al.* 1996; Bonhoeffer & Nowak 1997) and can be justified if the contribution of the immune responses to the death of infected cells is negligible or constant in time. The third assumption is that the free virus load is proportional to the infected cell load. This assumption is commonly made for HIV, because the dynamics of free virus are known to be much faster than those of the infected cells (Wei *et al.* 1995; Perelson *et al.* 1996; Ramratnam *et al.* 1999). For practical purposes, measurements of the free virus load of wt and mt virus can be used instead of the corresponding infected cell loads, assuming that the proportionality factor between free virus and infected cells does not change over the period of observation or the range of observed virus loads.

In this paper, we intentionally focused on drug resistance mutations in HIV. However, in principle, the approach is much more general. In particular, if the above assumptions can be justified for mutations in other viruses or affecting other parts of the HIV genome, then GM can be applied without modification. If, however, the mutations are known to affect the death rate of infected cells, GM needs to be modified accordingly. All methods discussed here assume competition between just two viral variants. However, in some cases, more than two predominant variants may grow simultaneously. This problem can be circumvented by restricting the analysis just to those periods in which only two variants predominate.

In summary, we have developed a method (GM), which reliably estimates the selection coefficient between two virus variants based on longitudinal data. Unlike the methods typically used in the literature, this method does not suffer from the often unrealistic assumption, that the replication rate of the virus population is constant in time. The estimated selection coefficient measures the replicative capacity of two competing variants relative to each other and thus is independent of factors that simultaneously affect the replication rate of both variants. Selection coefficients determined in this way should therefore be more meaningful for the comparison between experimental set-ups and between patients. A more accurate and reliable quantification of viral fitness will in turn improve our understanding of the evolutionary dynamics of drug resistance and may eventually help to establish a basis for a more rational prescription of effective therapy.

We thank G. A. Funk, V. Müller and S. W. D. Frost for discussion and comments. S.B. and A.D.B. gratefully acknowledge support from the Swiss National Science Foundation (grant nos: 631-062898 and 20-61753.00).

APPENDIX A

(a) *Expectations and variances*

Typically the experimental assays do not directly measure the log wt virus load (w_i) and the log mt/wt ratio (h_i), but instead measure the total virus load (V_i) and the fraction of mutant virus (f_i). Generally, the relationship between these quantities is given by $h_i = \ln[f_i/(1 - f_i)]$ and $w_i = \ln[(1 - f_i)V_i]$. However, as the experimental measurements of the total virus load and the fraction of mutants are subject to experimental error, the computation of the expected values of the log mt/wt ratio and the rescaled time need to take into account the experimental error in the original data. Using error propagation (Lynch & Walsh 1998) and ignoring all but first-order terms, one obtains the following expressions for the expected values and their variances

$$\hat{h}_i = \ln\left(\frac{\hat{f}_i}{1-\hat{f}_i}\right) + \frac{1}{2}\sigma_{f_i}^2 \frac{2\hat{f}_i - 1}{\hat{f}_i^2(1-\hat{f}_i)^2}, \tag{A 1}$$

$$\hat{\tau}_i = \ln\left(\frac{(1-\hat{f}_i)\hat{V}_i}{(1-\hat{f}_0)\hat{V}_0}\right) + \delta T_i + \frac{1}{2}\left(\frac{\sigma_{f_0}^2}{(1-\hat{f}_0)^2} - \frac{\sigma_{f_i}^2}{(1-\hat{f}_i)^2} + \frac{\sigma_{V_0}^2}{\hat{V}_i^2} - \frac{\sigma_{V_i}^2}{\hat{V}_i^2}\right), \tag{A 2}$$

$$\sigma_{h_i}^2 = \sigma_{f_i}^2 \frac{1}{\hat{f}_i^2(1-\hat{f}_i)^2}, \tag{A 3}$$

$$\sigma_{\tau_i}^2 = \frac{\sigma_{f_0}^2}{(1-\hat{f}_0)^2} + \frac{\sigma_{f_i}^2}{(1-\hat{f}_i)^2} + \frac{\sigma_{V_0}^2}{\hat{V}_i^2} + \frac{\sigma_{V_i}^2}{\hat{V}_i^2} + \sigma_\delta^2 T_i^2, \tag{A 4}$$

where $\hat{f}_i$ and $\hat{V}_i$ represent the experimental data (with corresponding variances $\sigma_{f_i}^2$ and $\sigma_{V_i}^2$). The variance in the denominator of equation (2.6) is given by $\mathrm{var}(h_0 + s\hat{\tau}_i - \hat{h}_i) = \sigma_{h_i}^2 + s^2\sigma_{\tau_i}^2 - 2s\,\mathrm{covar}(\hat{\tau}_i, \hat{h}_i)$, where $\mathrm{covar}(\hat{\tau}_i, \hat{h}_i) = \sigma_{f_i}^2/(f_i(f_i - 1)^2)$.

(b) *Estimation of infected cell death rates* in vitro

The appropriate procedure to estimate infected cell death rates *in vitro* clearly depends on the specific design of the growth competition assay. In principle, the death rate can be estimated by applying anti-retroviral drugs to a virus culture and by measuring the decline kinetics of infected cells. However, it is important to note that for growth competition assays based on serial transfer, the time between transfers has to be considerably larger than the expected lifespan of infected cells. Otherwise one needs to correct for the curtailed virus production due to the transfer of virus into fresh medium prior to the natural death of the infected cells.

REFERENCES

Bonhoeffer, S. & Nowak, M. A. 1997 Pre-existence and the emergence of drug resistance in HIV infection. *Proc. R. Soc. Lond.* B **264**, 631–637. (DOI 10.1098/rspb.1997.0089.)

Bonhoeffer, S., Coffin, J. M. & Nowak, M. A. 1997*a* HIV-drug therapy and virus load. *J. Virol.* **71**, 3725–3728.

Bonhoeffer, S., May, R. M., Shaw, G. M. & Nowak, M. A. 1997*b* Virus dynamics and drug treatment. *Proc. Natl Acad. Sci. USA* **94**, 6971–6976.

Borrow, P. (and 10 others) 1997 Antiviral pressure exerted by

HIV-1-specific cytotoxic T-lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat. Med.* **3**, 205–211.

Boucher, C. A. B., Tersmette, M., Lange, J. M. A., Kellam, P., Degoede, R. E. Y., Mulder, J. W., Darby, G., Goudsmit, J. & Larder, B. A. 1990 Zidovudine sensitivity of human immunodeficiency viruses from high-risk, symptom-free individuals during therapy. *Lancet* **336**, 585–590.

Chao, L. 1990 Fitness of RNA virus decreased by Muller's ratchet. *Nature* **348**, 454–455.

Croteau, G., Doyon, L., Thibeault, D., McKercher, G., Pilote, L. & Lamarre, D. 1997 Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors. *J. Virol.* **71**, 1089–1096.

Eastman, P. S. (and 10 others) 1998 Genotypic changes in human immunodeficiency virus type 1 associated with loss of suppression of plasma viral RNA levels in subjects treated with ritonavir (norvir) monotherapy. *J. Virol.* **72**, 5154–5164.

Goudsmit, J., De Ronde, A., Ho, D. D. & Perelson, A. S. 1996 Human immunodeficiency virus fitness *in vivo*: calculations based on a single zidovudine resistance mutation at codon 215 of reverse transcriptase. *J. Virol.* **70**, 5662–5664.

Goudsmit, J., De Ronde, A., De Rooij, E. & De Boer, R. 1997 Broad spectrum of *in vivo* fitness of human immunodeficiency virus type 1 subpopulations differing at reverse transcriptase codons 41 and 215. *J. Virol.* **71**, 4479–4484.

Goulder, P. J. (and 11 others) 1997 Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat. Med.* **3**, 212–217.

Goulder, P. J. (and 18 others) 2001 Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* **412**, 334–338.

Harrigan, P. R., Bloor, S. & Larder, B. A. 1998 Relative replicative fitness of zidovudine-resistant human immunodeficiency virus type 1 isolates *in vitro*. *J. Virol.* **72**, 3773–3778.

Ho, D. D., Toyoshima, T., Mo, H. M., Kempf, D. J., Norbeck, D., Chen, C. M., Wideburg, N. E., Burt, S. K., Erickson, J. W. & Singh, M. K. 1994 Characterization of human immunodeficiency virus type 1 variants with increased resistance to a c-2-symmetrical protease inhibitor. *J. Virol.* **68**, 2016–2020.

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. & Markowitz, M. 1995 Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126.

Holland, J. J., de la Torre, J. C., Clarke, D. K. & Duarte, E. 1991 Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *J. Virol.* **65**, 2960–2967.

Larder, B. A. & Kemp, S. D. 1989 Multiple mutations in HIV-1 reverse-transcriptase confer high-level resistance to zidovudine (AZT). *Science* **246**, 1155–1158.

Larder, B. A., Darby, G. & Richman, D. D. 1989 HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science* **243**, 1731–1734.

Lynch, M. & Walsh, B. 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer.

Mansky, L. M. 1996 Forward mutation rate of human immunodeficiency virus type 1 in a T-lymphoid cell line. *AIDS Res. Hum. Retroviruses* **12**, 307–314.

Marée, A. F., Keulen, W., Boucher, C. A. & De Boer, R. J. 2000 Estimating relative fitness in viral competition experiments. *J. Virol.* **74**, 11 067–11 072.

Martinez, M. A., Carrillo, C., Gonzalez-Candelas, F., Moya, A., Domingo, E. & Sobrino, F. 1991 Fitness alteration of foot-and-mouth disease virus mutants: measurement of adaptability of viral quasispecies. *J. Virol.* **65**, 3954–3957.

Martinez-Picado, J., Savara, A. V., Sutton, L. & D'Aquila, R. T. 1999 Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. *J. Virol.* **73**, 3744–3752.

Martinez-Picado, J., Savara, A. V., Shi, L., Sutton, L. & D'Aquila, R. T. 2000 Fitness of human immunodeficiency virus type 1 protease inhibitor-selected single mutants. *Virology* **275**, 318–322.

Nagylaki, T. 1992 *Introduction to theoretical population genetics*. Berlin: Springer.

Nowak, M. A., Bonhoeffer, S., Loveday, C., Balfe, P., Semple, M., Kaye, S., Tenantflowers, M. & Tedder, R. 1995 HIV results in the frame: results confirmed. *Nature* **375**, 193.

Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., Thomas, H. C. & McDade, H. 1996 Viral dynamics of HBV infection. *Proc. Natl Acad. Sci. USA* **93**, 4398–4402.

Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. 1996 HIV-1 dynamics *in vivo*: virion clearance rate, infected cell lifespan, and viral generation time. *Science* **271**, 1582–1586.

Phillips, R. E. (and 10 others) 1991 Human immunodeficiency virus genetic variation that can escape cytotoxic T-cell recognition. *Nature* **354**, 453–459.

Ramratnam, B., Bonhoeffer, S., Binley, J., Hurley, A., Zhang, L., Mittler, J. E., Markowitz, M., Moore, J. P., Perelson, A. S. & Ho, D. D. 1999 Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet* **354**, 1782–1785.

Ribeiro, R. M. & Bonhoeffer, S. 2000 Production of resistant HIV mutants during antiretroviral therapy. *Proc. Natl Acad. Sci. USA* **97**, 7681–7686.

Ribeiro, R. M., Bonhoeffer, S. & Nowak, M. A. 1998 The frequency of drug resistant mutants before therapy. *AIDS* **12**, 461–465.

Richman, D. D. 1990 Zidovudine resistance of human immunodeficiency virus. *Rev. Infect. Dis.* **12**(Suppl. 5), S507–S512.

Richman, D. D. (and 12 others) 1994 Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy. *J. Virol.* **68**, 1660–1666.

StClair, M. H., Martin, J. L., Tudorwilliams, G., Bach, M. C., Vavro, C. L., King, D. M., Kellam, P., Kemp, S. D. & Larder, B. A. 1991 Resistance to DDI and sensitivity to AZT induced by a mutation in HIV-1 reverse-transcriptase. *Science* **253**, 1557–1559.

Wei, X. (and 11 others) 1995 Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122.

Yuste, E., Sanchez-Palomino, S., Casado, C., Domingo, E. & Lopez-Galindez, C. 1999 Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *J. Virol.* **73**, 2745–2751.

Zennou, V., Mammano, F., Paulous, S., Mathez, D. & Clavel, F. 1998 Loss of viral fitness associated with multiple gag and gag-pol processing defects in human immunodeficiency virus type 1 variants selected for resistance to protease inhibitors *in vivo*. *J. Virol.* **72**, 3300–3306.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.