The frequency of resistant mutant virus before antiviral therapy

Ruy M. Ribeiro, Sebastian Bonhoeffer* and Martin A. Nowak

Objective: To calculate the expected prevalence of resistant HIV mutants before antiviral therapy.

Design: HIV replication generates virus mutants. The prevalence of these mutants is determined by mutation and selection/fitness. Some mutations will confer drug resistance and it is crucial for the success of antiviral drug therapy to determine whether these resistant viruses are present before the initiation of therapy.

Methods: A quasispecies equation was used to calculate the expected frequency of drug-resistant virus prior to therapy.

Results and conclusions: We show how the pretreatment frequency of resistant virus depends on the number of point mutations between wild-type and mutant virus, the selective disadvantage of the resistant mutant and the intermediate mutants, and the mutation rate.

AIDS 1998, **12**:461–465

Keywords: Antiviral therapy, mathematical model, resistant mutant, viral load

Introduction

A number of potent anti-HIV drugs that inhibit the replication of the virus *in vivo* are now available. Reverse transcriptase inhibitors prevent the infection of new cells, while protease inhibitors prevent HIV-infected cells from producing infectious virus particles. Drug treatment usually results in a rapid decline of plasma virus load and an increase in the CD4 cell count [1–5]. Monotherapy often leads to rapid emergence of drug-resistant virus mutants. For some drugs, a single point mutation can confer high level resistance, while for other drugs several point mutations are required. Combination therapy can result in a longer lasting suppression of virus load. Lamivudine and zidovudine together maintain an approximately 10-fold reduction of plasma virus load in patients treated for up to (approximately) 1 year [6]. Combining lamivudine, zidovudine and one (or two) protease inhibitors can reduce virus load by more than 10 000-fold [7,8]. Patients have undetectable plasma virus, declining levels of infectious HIV-1 in peripheral blood mononuclear cells and slowly declining amounts of HIV-1 DNA provirus.

The success of antiviral therapy depends (amongst other things) crucially on whether resistant mutant virus is present before the initiation of therapy [9,10]. If resistant virus is present prior to treatment, then application of the drug leads to declining levels of sensitive wild-type virus and increasing levels of resistant mutant virus. Mutant virus grows either because of an increased supply of target cells [11,12] or a reduction in antiviral immunity. If there is only a small probability that resistant virus is present in a patient before antiviral therapy, then a straightforward calculation shows that the chance that it will arise after the initiation of therapy is even smaller [13,14].

We call a virus mutant 'resistant' if it has a positive growth rate in the presence of therapy. For single drug therapy, it is likely that certain one or two point mutations confer resistance. Therefore, it seems likely that virus mutants, resistant to almost any single drug

From the Department of Zoology, University of Oxford, Oxford, UK and the *Aaron Diamond AIDS Research Center, New York, New York, USA.

Sponsorship: Supported by Praxis XXI from Junta Nacional de Investigação Científica e Tecnológica (R.M.R.) and the Wellcome Trust (S.B. and M.A.N.).

Requests for reprints to: Professor Martin Nowak, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Date of receipt: 28 May 1997; revised: 5 December 1997; accepted: 11 December 1997.

therapy, are present in a patient before treatment. Resistance to two or three drugs may require two or more point mutations.

Mathematical models of antiviral drug treatment have been developed to study the effect of increasing target cell abundance on rebound of wild-type virus [15], the emergence of resistant mutant virus as a consequence of increasing target cell abundance [11], the effect of treatment on reducing viral diversity [16], the effect on virus load of a drug-induced reduction of infectivity [17], and the emergence of several different mutants of increasing levels of resistance during therapy [18,19]. Mathematical models of antiviral therapy have also provided estimates for kinetic parameters of virus replication *in vivo* [1,2,4,7,20–23]. De Boer and Boucher developed a mathematical model to study the consequence of preventing CD4 cell increase during therapy on the emergence of resistant virus [24].

In this article we use a quasispecies equation of viral dynamics [25] to calculate the expected frequency of drug-resistant virus in untreated patients. We will provide results for the frequency of resistant mutants that differ in one, two, three or more point mutations from wild-type virus. We will compare the effect of mutation and selection on pretreatment frequency, and also study the consequence of various degrees of selection disadvantage of intermediate mutants.

The model

We begin with a simple model that contains wild-type virus and one mutant. It is assumed here that wild-type and mutant differ by only a single point mutation (Fig. 1a). The mutation rate is given by the parameter μ. Furthermore, it is assumed that wild-type virus has a fitness advantage over mutant virus in the absence of drug; the selection coefficient is given by s. The model has three variables: uninfected cells, x; cells productively infected by wild-type virus, y_0; and by mutant virus, y_1. Uninfected cells are produced at rate λ and die at rate dx. Infected cells die at rates ay_0 and ay_1. Infected cells give rise to new infected cells at a rate proportional to the abundance of uninfected cells multiplied by infected cells; the rate constant is β. It is assumed that free virus dynamics are fast compared with infected cell turnover [22]. Therefore, we do not have to consider a separate equation for free virions, but simply assume that virion abundance is proportional to infected cell abundance. The model equations are as follows:

$$\dot{x} = \lambda - dx - \beta x[y_0 + (1 - s)y_1]$$

$$\dot{y}_0 = \beta x[(1 - \mu)y_0 + (1 - s)\mu y_1] - ay_0$$

$$\dot{y}_1 = \beta x[\mu y_0 + (1 - s)(1 - \mu)y_1] - ay_1$$

At equilibrium, the ratio of mutant to wild-type virus, to first order in μ, is as follows:

$$y_1^*/y_0^* = \mu/s$$

For example, if the point mutation rate is $\mu = 3 \times 10^{-5}$ [26] and the mutant has a selective disadvantage of $s = 0.01$ [27], then the relative proportion of mutant to wild-type is 3×10^{-3}. In other words, for this choice of parameters, about 1 in 300 cells contains the mutant virus. If the mutant had a large selective disadvantage ($s \approx 1$), then the ratio of mutant to wild-type would be approximately 10^{-5}. However, there is some

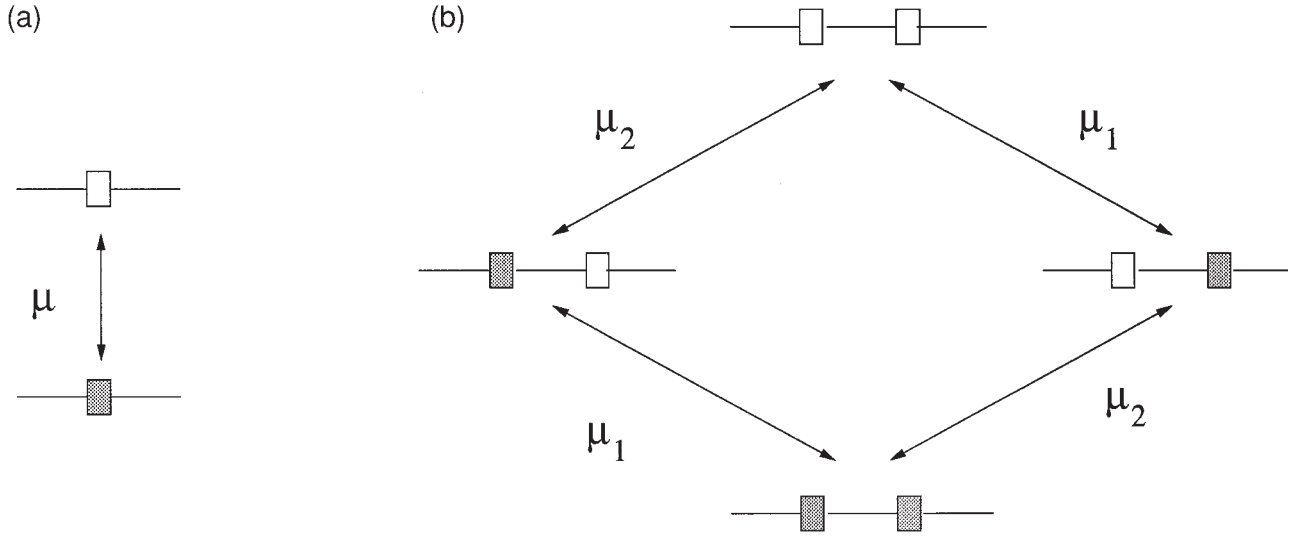

Fig. 1. A schematic illustration for the system with (a) one and (b) two point mutations between sensitive wild-type virus and resistant mutant virus. The empty box represents wild-type virus and the shaded box represents mutant virus.

experimental evidence for one-error mutants with frequencies of the order of 10^{-3} [28] in previously untreated patients.

We now expand the model to consider resistant mutants that differ by two point mutations (Fig. 1b). This model has five variables: uninfected cells, x; cells infected by wild-type virus, y_{00}; cells infected by one-error mutants, y_{01} and y_{10}; and cells infected by the two-error mutant, y_{11}. Note that we use a binary notation for wild-type virus (00) and the various mutants (01, 10, and 11). Thus, at each of two relevant positions we consider one amino-acid substitution. We assume that these amino-acid substitutions are due to a single base substitution in the viral genome. The replication fidelity of the reverse transcriptase may vary for different positions in the genome and therefore we assume that the mutation rate for the first position is μ_1 and for the second position is μ_2. The selective disadvantage of the mutants y_{01}, y_{10} and y_{11} is given by s_{01}, s_{10} and s_{11}, respectively. The model equations are as follows:

$$\dot{x} = \lambda - dx - \beta x(y_{00}+r_{01}y_{01}+r_{10}y_{10}+r_{11}y_{11})$$

$$\dot{y}_{00}=\beta x(\nu_1\nu_2 y_{00}+r_{01}\nu_1\mu_2 y_{01}+r_{10}\nu_2\mu_1 y_{10}+r_{11}\mu_1\mu_2 y_{11}) - a y_{00}$$

$$\dot{y}_{01}=\beta x(\nu_1\mu_2 y_{00}+r_{01}\nu_1\nu_2 y_{01}+r_{10}\mu_1\mu_2 y_{10}+r_{11}\mu_1\nu_2 y_{11}) - a y_{01}$$

$$\dot{y}_{10}=\beta x(\mu_1\nu_2 y_{00}+r_{01}\mu_1\mu_2 y_{01}+r_{10}\nu_1\nu_2 y_{10}+r_{11}\nu_1\mu_2 y_{11}) - a y_{10}$$

$$\dot{y}_{11}=\beta x(\mu_1\mu_2 y_{00}+r_{01}\mu_1\nu_2 y_{01}+r_{10}\nu_1\mu_2 y_{10}+r_{11}\nu_1\nu_2 y_{11}) - a y_{11}$$

The abbreviations $\nu_i = 1 - \mu_i$ and $r_{ij} = 1 - s_{ij}$ have been used. The exact equilibrium solution of this system is complicated, but we can obtain an elegant approximation by neglecting back-mutations. In this case, we simply solve the following eigenvalue equation:

$$\begin{bmatrix} \nu_1\nu_2 & 0 & 0 & 0 \\ \nu_1\mu_2 & r_{01}\nu_1 & 0 & 0 \\ \nu_2\mu_1 & 0 & r_{10}\nu_2 & 0 \\ \mu_1\mu_2 & r_{01}\mu_1 & r_{10}\mu_2 & r_{11} \end{bmatrix} \times \begin{bmatrix} y_{00} \\ y_{01} \\ y_{10} \\ y_{11} \end{bmatrix} = \wedge \times \begin{bmatrix} y_{00} \\ y_{01} \\ y_{10} \\ y_{11} \end{bmatrix}$$

We obtain the following:

$$y^\star_{01} / y^\star_{00} = \frac{\mu_2}{s_{01} - \mu_2}$$

$$y^\star_{10} / y^\star_{00} = \frac{\mu_1}{s_{10} - \mu_1}$$

$$y^\star_{11} / y^\star_{00} = \frac{\mu_1\mu_2(\mu_1\mu_2 - \mu_1 - \mu_2 + s_{01} + s_{10} - s_{01}s_{10})}{(s_{01} - \mu_2)(s_{10} - \mu_1)(\mu_1\mu_2 - \mu_1 - \mu_2 + s_{11})}$$

If in addition we assume that the selection coefficients are larger than the mutation rates, then we obtain the following approximate equilibrium ratios:

$$y^\star_{01} / y^\star_{00} = \mu_2 / s_{01}$$

$$y^\star_{10} / y^\star_{00} = \mu_1 / s_{10}$$

$$y^\star_{11} / y^\star_{00} = (\mu_1\mu_2 / s_{11})\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1\right)$$

Again, this approximation is accurate if the selection coefficients are noticeably larger than the mutation rates. For example, if $\mu_1 = \mu_2 = 3 \times 10^{-5}$ and if $s_{01} = s_{10} = s_{11} = 0.01$ then $y_{11}/y_{00} \approx 2 \times 10^{-5}$. In other words, approximately 1 in 50 000 productively infected cells contains resistant mutant virus, which differs from the wild-type by two point mutations. Table 1 shows the effect of different selective disadvantages for the intermediate mutants, s_{01} and s_{10}, on the relative frequency of y_{11}. For example, if $s_{11} = 0.01$, but $s_{01} = s_{10} = 1$ (i.e., the double mutant has a small selective disadvantage, but the intermediate one-error mutants cannot replicate) then the relative abundance of the double mutant is $y_{11}/y_{00} \approx 9 \times 10^{-8}$. If, on the other hand, $s_{01} = s_{10} = 0.001$ (i.e., the intermediate mutants have a selective disadvantage of only 0.1% compared with wild-type) then we find $y_{11}/y_{00} \approx 1.8 \times 10^{-4}$. The difference in the equilibrium frequency of the double mutant is several orders of magnitude. Therefore, in order to estimate the pretreatment equilibrium frequency of a resistant variant, it is, of course, not sufficient to know the selective disadvantage of this mutant compared with wild-type, but it is also necessary to know the selection coefficients of all intermediate variants.

The above model can be expanded to consider mutants that differ from wild-type virus by more than two point mutations. For simplicity, assume that the average mutation rate for each site is the same, μ, and that all mutants have the same selective disadvantage, s, which is much smaller than 1. In this case we obtain (Fig. 2) the following:

for one-error mutants, $y^\star_1/y^\star_0 = \mu/s$;

for two-error mutants, $y^\star_{11}/y^\star_{00} = 2(\mu/s)^2$;

for three-error mutants, $y^\star_{111}/y^\star_{000} = 6(\mu/s)^3$;

for n-error mutants, $y^\star_{11..1}/y^\star_{00..0} = n!(\mu/s)^n$.

Table 1. Effect of the selective disadvantages of the intermediate one-error mutants, y_{01} and y_{10}, on the relative abundance of the two-error mutant, y_{11}.

s_{01},s_{10}	10^{-3}	10^{-2}	10^{-1}	1
10^{-3}	1.8×10^{-4}	9.9×10^{-5}	9.1×10^{-5}	9.0×10^{-5}
10^{-2}		1.8×10^{-5}	9.9×10^{-6}	9.0×10^{-6}
10^{-1}			1.7×10^{-6}	9.0×10^{-7}
1				9.0×10^{-8}

A mutation rate of $\mu = 3 \times 10^{-5}$ and a selective disadvantage for y_{11} of $s_{11} = 0.01$ are assumed. The selective disadvantages, s_{01} and s_{10}, of the intermediate mutants are varied between 0.001 and 1. The equilibrium frequency, relative to wild-type, of the resistant y_{11} mutant is shown. (Since the effect of s_{01} and s_{10} is the same, only half of the symmetric table is presented, for simplicity.)

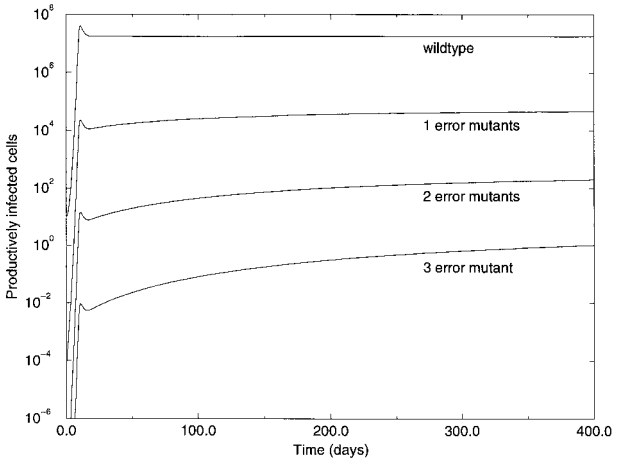

Fig. 2. Time evolution of a system with three point mutations. Originally the patient is infected by wild-type virus. The increase in total abundance of wild-type virus, one-, two-, and three-error mutants is shown. Parameter values are as follows: $\beta = 5 \times 10^{-8}$, $s_i = 0.01$, $\lambda = 10^7$, $a = 0.5$, $d = 0.1$ and $\mu_i = 3 \times 10^{-5}$.

Of course, $n!(\mu/s)^n$ has to be less than unity otherwise our approximations break down. This is essentially an error threshold condition of quasispecies theory [29,30].

Table 2 shows equilibrium frequencies of mutants that differ from wild-type by up to five point mutations. For $s = 0.01$, the frequency of a three-error mutant is 1.6×10^{-7}, of a four-error mutant is 1.9×10^{-9} and of a five-error mutant is 2.9×10^{-11}. Suppose that there are approximately 10^7–10^8 [31] productively infected cells in an HIV-1-infected patient. For $s = 0.01$, a particular four- or five-error mutant is unlikely to be present, while the three-error mutant is likely to exist. For a higher selective disadvantage ($s > 0.01$) the three-error mutant is also unlikely to exist in a patient prior to treatment. We can also ask what is the maximum selective disadvantage that is still compatible with the survival of a given virus mutant in a patient with N infected cells. If $N = 10^8$ and $\mu = 3 \times 10^{-5}$, then a two-error mutant will survive if $s < 0.4$, a three-error mutant if $s < 0.03$, and a four-error mutant if $s < 0.007$. (In this calculation, intermediate mutants have the same selective disadvantage, s. Other assumptions are also possible and can be calculated.)

Table 2. Equilibrium frequencies of mutants that differ from the wild-type by one, two, three, four or five point mutations.

	s		
n	0.001	0.01	0.1
1	3.0×10^{-2}	3.0×10^{-3}	3.0×10^{-4}
2	1.8×10^{-3}	1.8×10^{-5}	1.8×10^{-7}
3	1.6×10^{-4}	1.6×10^{-7}	1.6×10^{-10}
4	1.9×10^{-5}	1.9×10^{-9}	1.9×10^{-13}
5	2.9×10^{-6}	2.9×10^{-11}	2.9×10^{-15}

For this example, it is assumed that all intermediate mutants have (essentially) the same selective disadvantage, s (range, 0.001–0.1). The mutation rate is $\mu = 3 \times 10^{-5}$.

The model also provides insight into the time evolution of resistant mutant virus during the course of infection. If a patient is initially infected by sensitive wild-type virus, y_0, then the frequency of resistant mutant virus, y_1, rises as follows:

$$y_1(t)/y_0(t) = (1 - \exp^{-\beta x(s-\mu)t})\mu/(s-\mu)$$

Note that this calculation assumes that the number of uninfected cells, x, is approximately constant during the rise of resistant mutants. The mutant will arrive at half its equilibrium value after $(\ln 2)/[\beta x(s-\mu)]$ days. In principle, we can recursively obtain the rate of convergence towards equilibrium for any of the mutant strains, but the analytical expressions are too complex to be shown here. However, it is interesting to compare the times until a two-error (or greater) mutant attains a given frequency relative to the wild-type if (i) the intermediate mutants cannot replicate at all, or (ii) all mutants have the same selective disadvantage. For example, a two-error mutant reaches a frequency of 10^{-8} about five times faster under scenario (ii) than under scenario (i). Hence, the presence of intermediate mutants with positive growth rate not only strongly affects the equilibrium frequency of a n-error mutant, but also its rate of ascent. Note, however, that the time necessary for a mutant to reach a given percentage of its equilibrium value is larger under scenario (ii), simply because the equilibrium value is much higher in this scenario.

Conclusions

The success of drug therapy depends to a large extent on whether resistant virus is present in patients prior to treatment. Based on a quasispecies model for the balance of selection and mutation between wild-type and mutants, we derived analytical expressions for the pre-treatment frequency of mutant virus. We showed that the mutant frequency depends on the number of point mutations between wild-type and mutant, the selective disadvantage of the mutant and all its intermediates, and the mutation rate. For example, for a mutation rate of 3×10^{-5} [26] and a selective disadvantage for the mutant and all intermediates of 1%, we find that the frequency of a three-error mutant relative to wild-type is approximately 10^{-7}. If there are on average 10^7–10^8 productively infected cells in a patient [31], then such a mutant may well be present before the start of therapy.

There are several practical conclusions to be drawn. The probability of treatment failure due to viral resis-

tance can be reduced in three ways: (i) treatment should start early when virus load is still low and the frequency of resistant mutants is small (assuming that most patients are infected by drug-sensitive wild-type virus); (ii) treatment should commence immediately with multiple drugs (three or more) – it is clearly disadvantageous to initiate therapy with one drug and then add other drugs later, giving the virus the possibility to develop resistance to each drug sequentially; and (iii) treatment should combine drugs for which resistance mutations are known to involve a considerable selective disadvantage in the absence of drug. It may be of great importance to identify such combinations, since mutants resistant to these combinations may not exist in patients prior to treatment.

References

1. Wei X, Ghosh SK, Taylor ME, *et al.*: **Viral dynamics in human immunodeficiency virus type 1 infection.** *Nature* 1995, **373**:117–122.
2. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M: **Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.** *Nature* 1995, **373**:123–126.
3. Schuurman R, Nijhuis M, van Leeuwen R, *et al.*: **Rapid changes in human immunodeficiency virus type 1 RNA load and appearance of drug-resistant virus populations in persons treated with lamivudine (3TC).** *J Infect Dis* 1995, **171**:1411–1419.
4. Nowak MA, Bonhoeffer S, Loveday C, *et al.*: **HIV-1 dynamics: results confirmed.** *Nature* 1995, **375**:193.
5. Loveday C, Kay S, Tenant-Flowers M, *et al.*: **HIV-1 RNA serum-load and resistant viral genotypes during early zidovudine therapy.** *Lancet* 1995, **345**:820–824.
6. Eron JJ, Benoit SL, MacArthur RD, *et al.*: **Treatment with lamivudine, zidovudine, or both in HIV-positive patients with 200–500 CD4+ cells per cubic millimeter.** *N Engl J Med* 1995, **333**:1662–1669.
7. Perelson AS, Essunger P, Cao Y, *et al.*: **Decay characteristics of HIV-1-infected compartments during combination therapy.** *Nature* 1997, **387**:188–191.
8. Lafeuillade A, Poggi C, Sayada C, Pellegrino P, Profizi N: **Focusing on the second phase of plasma HIV-1 RNA clearance [letter].** *AIDS* 1997, **11**:264–266.
9. Coffin JM: **HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy.** *Science* 1995, **267**:483–489.
10. Coffin JM: **HIV viral dynamics.** *AIDS* 1996, **10** (suppl 3):S75–S84.
11. McLean AR, Nowak MA: **Competition between zidovudine-sensitive and zidovudine-resistant strains of HIV.** *AIDS* 1992, **6**:71–79.
12. Nowak MA, May RM: **AIDS pathogenesis: mathematical models of HIV and SIV infections.** *AIDS* 1993, **7** (suppl 1):S3–S18.
13. Bonhoeffer S, Nowak MA: **Pre-existence and emergence of drug resistance in HIV-1 infection.** *Proc R Soc Lond B Biol Sci* 1997, **264**:631–637.
14. Bonhoeffer S, May RM, Shaw GM, Nowak MA: **Models of virus dynamics and drug therapy.** *Proc Natl Acad Sci USA* 1997, **94**:6971–6976.
15. McLean AR, Emery VC, Webster A, Griffiths PD: **Population dynamics of HIV within an individual after treatment with zidovudine.** *AIDS* 1991, **5**:485–489.
16. Nowak MA, Anderson RM, McLean AR, Wolfs T, Goudsmit J, May RM: **Antigenic diversity thresholds and the development of AIDS.** *Science* 1991, **254**:963–969.
17. Bonhoeffer S, Coffin JM, Nowak MA: **Human immunodeficiency virus drug therapy and virus load.** *J Virol* 1997, **71**:3275–3278.
18. Frost SDW, McLean AR: **Quasispecies dynamics and the emergence of drug resistance during zidovudine therapy of HIV infection.** *AIDS* 1994, **8**:323–332.
19. Stilianakis NI, Boucher CAB, De Jong MD, van Leeuwen R, Schuurman R, De Boer RJ: **Clinical data sets of human immunodeficiency virus type-1 reverse transcriptase resistant mutants explained by a mathematical model.** *J Virol* 1997, **71**:161–168.
20. Nowak MA, Anderson RM, Boerlijst MC, Bonhoeffer S, May RM, McMichael AJ: **HIV evolution and disease progression.** *Science* 1996, **274**:1008–1010.
21. Nowak MA, Bonhoeffer S, Shaw GM, May RM: **Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations.** *J Theor Biol* 1997, **184**:203–217.
22. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD: **HIV-1 dynamics *in vivo*: virion clearance rate, infected cell lifespan, and viral generation time.** *Science* 1996, **271**:1582–1585.
23. Herz AVM, Bonhoeffer S, Anderson RM, May RM, Nowak MA: **Viral dynamics *in vivo*: limitations on estimates of intracellular delay and virus decay.** *Proc Natl Acad Sci USA* 1996, **93**:7247–7251.
24. De Boer RC, Boucher CAB: **Anti-CD4 therapy of AIDS suggested by mathematical models.** *Proc R Soc Lond B Biol Sci* 1996, **263**:899–905.
25. Boerlijst MC, Bonhoeffer S, Nowak MA: **Viral quasi-species and recombination.** *Proc R Soc Lond B Biol Sci* 1996, **263**:1577–1584.
26. Mansky LM, Temin HM: **Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than the predicted from the fidelity of purified reverse transcriptase.** *J Virol* 1995, **29**:5087–5094.
27. Goudsmit J, De Ronde A, Ho DD, Perelson AS: **Human immunodeficiency virus fitness *in vivo*: calculations based on a single zidovudine resistance mutation at codon 215 of reverse transcriptase.** *J Virol* 1996, **70**:5662–5664.
28. Nájera I, Holguín A, Quiñones-Mateu ME, *et al.*: ***pol* gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy.** *J Virol* 1995, **69**:23–31.
29. Eigen M, Schuster P: **The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle.** *Naturwissenschaften* 1977, **64**:541–565.
30. Nowak MA, Schuster P: **Error thresholds of replication in finite populations. Mutation frequencies and the onset of Muller's ratchet.** *J Theor Biol* 1989, **137**:375–395.
31. Chun TW, Carruth L, Finzi D, *et al.*: **Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection.** *Nature* 1997, **387**:183–188.