HIV-1: Gambling on the evolution of drug resistance?

Despite the huge size of the HIV population in an infected patient, chance has an unexpected influence on its evolution.

ANDREW J. LEIGH BROWN[1] & DOUGLAS D. RICHMAN[2]

The increasing appreciation of the massive replication rates of human immunodeficiency virus (HIV) and the virus's high mutation rate gives the impression of a population of genetic variants of almost unlimited size that can adapt to any selective pressure that arises. Although there is much truth to this, a number of clinical observations clearly demonstrate that constraints on the evolution of HIV populations exist, and that chance often plays a significant role in disease course and response to treatment. The role of chance dictates reconsideration of the appropriate population genetic model to be applied to HIV, and has implications for both the natural history and the treatment of HIV disease.

Infection with HIV is characterized by high levels of virus in blood and lymphoid tissue with remarkably rapid turnover rates of actively replicating virus[1]. High concentrations of viral nucleic acid have been demonstrated in lymphoid tissue[2,3], and most infected adults have between 10^3 and 10^6 copies of plasma HIV RNA (in virions) per milliliter of plasma[4]. These virions have a maximal half-life of approximately 6 hours[5]. Because HIV's replication systems error prone and lacks any proof-reading mechanism[6,7] (like other single-stranded RNA viruses), the virus has a very high mutation rate — approximately 3×10^{-5} nucleotides per replication cycle[8]. Most primary infections with HIV appear to be monoclonal, with relatively homogeneous viral nucleotide sequences in newly infected patients. However, because of its high rates of replication and genetic variation, HIV undergoes rapid and extensive evolution within each infected patient during the course of infection[9,10].

It has been effectively argued that the high rate of virus replication drives viral evolution[11]. It is estimated that about 140 replication cycles (generations) occur annually, with a daily production of about 10^{10} virions[5]. Because of these estimates, the viral population has usually been considered effectively infinite. Consequently, the evolution of the quasi-species has been thought to be determined solely by selective pressures[11]. However, replication of HIV — and thus the response of the viral population to selection — also depends on the availability of susceptible $CD4^+$ T cells. Thus the delay in emergence of wild type virus when therapy is stopped is expected[12]. Furthermore, viral populations are not theoretical concepts embodied in deterministic models but are real biologic populations, and chance events can have a profound effect on evolutionary pathways. When an advantageous mutation first occurs, it is more likely to be lost than to spread. The probability of eventual success in a real population (formally known as *fixation*) is approximately equal to twice the fitness advantage. Thus for every mutant with a five percent advantage that is successful, nine others arise but are lost[13,14].

The effect of chance is even more pronounced in small or subdivided populations. In this case the genetic structuring of the population is summarized by the term "effective population number" (N_e, see Box 1). The value of N_e is inversely related to the probability that the copies of the same gene in two randomly chosen genomes were derived from the same ancestral copy[14,15]. In a truly infinite population, this probability is zero by definition, and the effective size is infinite. Several factors can reduce N_e below N, including subdivision of the population, fluctuations in population size and even natural selection. All these processes increase the chance that any two viral genomes share a common ancestry, and most of these processes can expected to apply to HIV.

The ratio of N_e and N determines the balance between the forces of selection and random genetic drift. The smaller the value of N_e, the less effective selection can be and the larger the role that chance may play in the evolution of the population. When N_e is low, even deleterious alleles will occasionally be successful. A knowledge of the value of N_e for HIV is therefore important for assessing the impact of fitness differences on the evolutionary fate of the viral population. It is also important in indicating whether the deterministic model employed to date or a stochastic model (in which chance effects are permitted) is the most appropriate population model for describing HIV populations (Box 1).

Box 1 Stochastic and deterministic models

Deterministic models deal with a theoretical population whose size is assumed to be unlimited (infinite). Consequently, the relative fitness of different allelic variants directly determines their abundance among the progeny virus after replication. Fitness differences of five or ten percent are thus translated into a five or ten percent frequency change from one replication cycle to the next. Stochastic models recognize that populations are finite. In this case a fitness difference is a difference in probability; that is, a five percent difference indicates that the mean difference over a large number of replicate populations will be five percent, but in any individual population chance events could increase, decrease or even reverse this difference. The impact of sampling is illustrated by a simple analogy — if a coin is tossed 100 times or more, the number of heads will approximate to 50, but if the record for every 10 coin tosses is examined, these are likely to range from 3/10 to 7/10. In the evolution of natural populations, therefore, the importance of chance is inversely related to the value of N_e.

Estimation of N_e

N_e can be estimated from nucleotide diversity data[16]. The expected average number of pairwise nucleotide differences for haploid populations is equal to $2N_e\mu$, where μ is the mutation rate to selectively neutral substitutions[17]. In addition, the expectation of the time since the most recent common ancestor ("coalescence time") for lineages under neutral evolution is also $2N_e$ generations[16]. So data on both the mean time to a common ancestor, or the level of selectively neutral variation may be used to estimate N_e for HIV-1, given information on the *in vivo* mutation rate.

Evolution of drug resistance in HIV

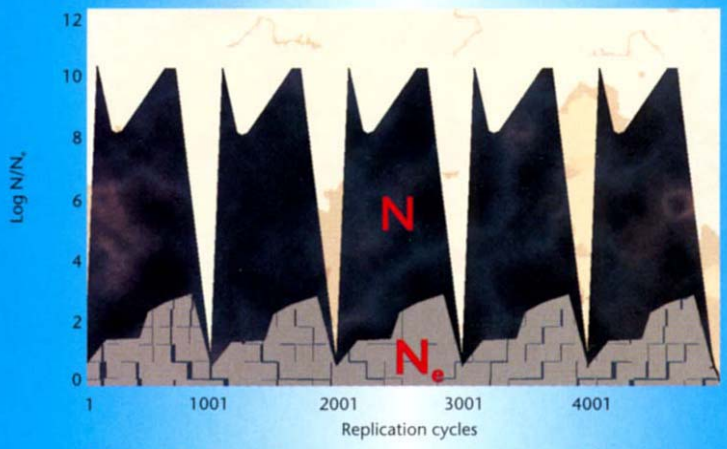
"Simple" resistance evolves quasi-deterministically. If HIV evolution were deterministic, the imposition of a selective pressure would rapidly select for the replacement of the common genotype by a single fitter genotype emerging from a pre-existing minority population that is constantly being generated as a conse-

quence of the high levels of virus replication. This is what appears to happen within weeks of starting treatment with the nucleoside analogue drug lamivudine (3TC): all circulating viral RNA contains valine at reverse transcriptase residue 184, a mutation that confers high-level drug resistance (compared to the wild type methionine) and a fitness advantage in the presence of drug of over 90 percent[18]. The 50 percent inhibitory concentration of lamivudine is increased over 1,000 fold with this single nucleotide change.

A single nucleotide change also confers high level resistance (100 to 1,000 fold reduction in susceptibility) to the non-nucleoside reverse transcriptase inhibitor nevirapine[19]. These mutants preexist in the plasma HIV RNA in a proportion of .001 to 1 percent of wild type[20]. The initiation of treatment selects for high-level resistance within weeks; however, individual patients display many different patterns of mutations[21]. Some patients acquire only a tyrosine to cysteine mutation at residue 181, which confers a level of resistance that permits residual antiviral activity in some patients. In other patients, nevirapine selects for an additional mutation at residue 108 that completely abrogates antiviral activity[22]. It is unclear whether this variation between individuals is attributable to chance or to mutational constraints conferred by the variable background sequences among patients.

Box 2 Effect of population fluctuation on effective population size

Changing population size has a dramatic effect on N_e, because variation is lost at each bottleneck. When population size fluctuates the maximum value for N_e has been shown to be equal to the harmonic mean of the population size in each generation[15]. On the basis of estimates of HIV population size and growth rate shortly after infection, we can describe the course of a "typical" HIV infection numerically, assuming infection is initiated by one or a few replication-competent virions and that the population may be expected to reach the order of 10^{10} particles by six weeks from the estimate of the length of the viral replication cycle obtained by Perelson *et al.*[3]. We may express the initial growth rate as about 0.5 log/generation. Under these conditions, the effective population size is less than 50 when N is 10^{10}. Extending the calculation to 1000 replication cycles (approximately seven years) assuming the usual sharp decrease in viral load at seroconversion of around three logs is followed by a gradual increase over the following years (Fig. 1), would predict a value of N_e of about 2×10^3 at 1000 replication cycles which is very close to what has been estimated[40]. If transmissions occurred every 1000 cycles, the relationship between N and N_e for five sequential infections would be as shown in Fig. 1. If transmissions are more frequent, N_e will be lower still.



Schematic depiction of the relationship between viral population number (viral load, N, plotted as $\log_{10}$) and effective population number (N_e) with time over five transmissions. For clarity, time is plotted as $\log_{10}$ replication cycle number for each in section separately. New infections are initiated after 1000 replication cycles (approximately seven years). More frequent transmissions would reduce N_e below the projected values.

Stochastic evolution of complex resistance patterns

The appearance of resistance mutations to zidovudine (AZT) also varies greatly among patients. Three to five different mutations must accumulate to confer a greater than 100-fold reduction in susceptibility to AZT[23]. Despite continual high-level virus replication in the presence of AZT, some patients develop detectable mutations within months, while others take the drug for years without evidence of emerging mutations[24]. Moreover, individuals display variability in the combinations of mutations that have been associated with AZT resistance[25].

Shirasaka *et al.* first described a distinctive pattern of high level, multiple nucleoside resistance[26]. Typically, a patient treated with AZT and didanosine (ddI) either sequentially or in combination will develop virus exhibiting more than 100-fold reductions in susceptibility not only to AZT and ddI, but to zalcitabine (ddC) and stavudine (d4T) as well. Virus from these patients first acquires a glutamine to methionine mutation at residue 151, and then acquires additional mutations at residues 62, 75, 77 and 116 (refs 27,28). Of particular interest, patients who first acquire a typical AZT resistance mutation, usually at residue 70, do not develop the 151 resistance complex. Conversely patients with the 151 resistance complex do not appear to acquire AZT resistance mutations. Virus from perhaps five to ten percent of HIV-infected patients follow this evolutionary pathway. Thus, after months to years, patients treated with AZT and ddI follow one of two mutually exclusive evolutionary pathways, apparently by chance.

The introduction of potent protease inhibitors provided several new observations regarding the highly variable and apparently random differences in the development of resistance. Monotherapy with drugs like indinavir or ritonavir reduce the level of plasma HIV RNA by approximately 99 percent[29,30]. Most patients over a period of 3 to 12 months, especially those treated with suboptimal doses, will experience a loss of suppression of virus replication that is attributable to the development of high level resistance (approximately 100-fold reductions in susceptibility). This resistance results from the cumulative acquisition of 4 to 7 mutations in the gene for protease[29,31]. These mutations occur in any of approximately 12 different residues in this 99-amino acid peptide. Almost no two isolates from treated patients have been shown to contain the same combination of mutations[29,31].

Approximately 30 percent of patients given the highest tolerated doses of the new potent protease inhibitors as monotherapy will experience sustained suppression of detectable HIV RNA or infectivity in blood for up to two years or more[30]. Upon drug withdrawal, no new mutations in the protease gene have been seen, even though several are known to

be found singly as polymorphisms in untreated patients[32,33]. No obvious characteristics distinguish those patients who will experience sustained suppression while complying with protease monotherapy from those who lose suppression and develop high level resistance. It is notable that the addition of nucleoside analogues to therapy with protease inhibitors can significantly increase the proportion with sustained suppression[30]. These observations suggest that — by chance — a minority of patients administered a potent protease inhibitor will experience complete suppression of virus replication without outgrowth of a highly resistant population, because unlike lamivudine or nevirapine, this requires the presence of multiple mutations. Moreover, the addition of nucleosides reduces the probability that a highly protease-resistant population of virus will grow out.

The syncytium-inducing (SI) phenotype

The appearance in some patients of virus with the syncytium-inducing (SI) phenotype provides another example of apparently stochastic evolution. Virus with the SI phenotype can be isolated from less than five percent of patients with a $CD4^+$ T lymphocyte count greater than 500 per microliter of blood[34]. During the course of HIV infection patients acquire SI virus in an apparently random manner. The proportion increases exponentially to 50 percent as the $CD4^+$ T cell count drops below 50 and patients die with AIDS. Nevertheless, half of the patients die without acquiring SI virus[34]. Those who do acquire the SI phenotype experience a three-fold more rapid rate of decline of CD4+ T lymphocytes and a shorter disease course[35,36], but the appearance of SI virus is stochastic. This evolutionary pathway, which requires the acquisition of several mutations in the HIV coat protein gp120, is followed apparently by chance in some patients but not others.

nef deletion mutants of HIV

During transmission with a low-infectious inoculum, a "genetic bottleneck" occurs at which N_e is close to 1, as evidenced by the usually homogenous initial population. When N_e is this low, "founder effects" can cause deleterious mutants to reach a high frequency[37]. This phenomenon has been seen in a cluster of patients found to be infected with virus bearing the same deletion mutant in the *nef* gene[38]. This highly attenuated virus is associated in these patients with very slow progression of disease[38,39].

Effective population for HIV *in vivo*

Using a newly developed algorithm, one of us has estimated the effective population number (N_e) of the viral population from nucleotide sequences of the V3 region from an infected hemophiliac with a total virus burden of 10^9 to be between 1.0 and 2.1×10^3 (ref. 40). More recently, this analysis has been extended to samples from three other asymptomatic patients. Each of these analyses generated estimates of N_e between 10^2 and 10^3 (A.J.L.B., L. Zhang & A. Mayer, manuscript in preparation). These values are substantially lower than might have been expected from the very large titer of plasma virus and low enough to indicate a significant role for chance in the evolution of the viral population.

Why is the estimate of N_e so far below the census size? Many factors act together to reduce N_e below N. First, estimates of the titer of infectious virus have been much lower than the titer of viral particles[4]. A proportion of newly replicated viral genomes are expected to carry an inactivating mutation that will prevent their replication. There is also likely to be substantial variation in the number of progeny particles from different infected cells. Some infected cells may die or be killed early after expression of viral antigens and before the successful generation of infectious progeny. A minority of infected cells may successfully generate many progeny that are genetically similar, and thus reduce the overall value of N_e substantially. Moreover, the successful propagation of progeny may be dependent upon geographical relationships of infected and susceptible cells, and not represent the genetic complexity of all infected cells on a one-for-one basis. The progeny of some cells may be poorly situated to infect susceptible host cells successfully before these virions are rapidly cleared, thus making propagation of that genotype less probable. On the other hand, another cell may generate genotypically identical progeny at a site where the opportunity to infect susceptible cells is high. This may be exemplified by the presence of focally distinctive genotypes in the spleen[41].

For all these reasons, the daily generation of 10^{10} virions may reflect an effective population size of 10^3. In addition, another significant factor reducing N_e within a patient is the extreme bottleneck in numbers that occurs successively as the virus is transmitted between individuals. This has an effect on N_e that is particularly strong for the first two or three years of the infection. During this time the diversity of the viral population remains low, because substantial time is required for the frequencies of new mutants to reach their equilibrium values (see Box 2).

What are the implications of the low effective number (N_e) of HIV? First, it clearly indicates that stochastic, rather than deterministic, models are more appropriate for describing the evolution of the viral population[42,43]. Second, the value of N_e sets a lower limit to the selective advantage required to determine the evolution of the viral population. For fitness differences less than the value of $1/N_e$, genetic drift is the primary force determining the fate of a genetic variant. When a mutation confers a fitness advantage of 50 percent or more, the outcome will be unaffected. However when fitness advantages are small, or when two mutations each confer a large, but similar fitness advantage, chance could play an important role, as the first mutation to arise will be likely to dominate the population even though others may be more fit.

The low effective number also provides an explanation for apparently chance differences in similar patients. Some develop SI variants; some do not. Some patients develop high level resistance to zidovudine and protease inhibitors and some do not. Individual patients develop very different patterns of resistance mutations to nucleosides, non-nucleoside reverse transcriptase inhibitors and protease inhibitors.

The lower effective population size has implications for chemotherapy. It explains why regimens with only three potent drugs can suppress virus replication in most patients to levels at which there is no evidence of ongoing replication despite initiating treatment at plasma levels of HIV RNA between 10^4 and 10^6 copies per ml (and many orders of magnitude higher levels in lymphoid tissue). Moreover, the rationale of treating early in infection is reinforced by the estimation of lower effective population numbers (and thus genetic complexity) at early stages (see Box 2). Virological studies in HIV therapeutics have been advanced substantially by the adoption of quantitative approaches to population dynamics. Incorporation of more accurate genetic models of the viral population will add even greater power to the analysis of antiviral resistance.

Acknowledgments

This work was supported by grants AI 27670, AI 38858, UCSD Center for AIDS Research (AI 36214), AI 29164, TW 00680 (Fogarty Award) from the National Institutes of Health; the Research Center for AIDS and HIV Infection of the San Diego Veterans Affairs Medical Center; and grants from the Medical Research Council.

1. Havlir, D.V. & Richman, D.D. Viral dynamics of HIV: Implications for drug development and therapeutic strategies. *Ann. Intern. Med.* **124**, 984–994 (1996).
2. Embretson, J. *et al.* Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature* **363**, 359–362 (1993).
3. Pantaleo, G. *et al.* HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature* **362**, 355–358 (1993).
4. Piatak, M., Jr. *et al.* High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science* **259**, 1749–1754 (1993).
5. Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M. & Ho, D.D. HIV-1 dynamics in vivo: Virion clearance rate, infected cell lifetime, and viral generation time. *Science* **271**, 1582–1586 (1996).
6. Eigen, M. New concepts for dealing with the evolution of nucleic acids. *Cold Spring Harbor Symposia on Quantitative Biology* **LII**, 307–320 (1987).
7. Drake, J.W. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90**, 4171–4175 (1993).
8. Mansky, L.M. & Temin, H.M. Lower in vivo mutation rate of human immunodeficiency virus-type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**, 5087–5094 (1995).
9. Wain-Hobson, S. The fastest genome evolution ever described: HIV variation in situ. *Curr. Op. Genet. Devel.* **3**, 878–883 (1993).
10. Leigh-Brown, A.J. & Holmes, E.C. The evolutionary biology of human immunodeficiency virus. *Ann. Rev. Ecol. Sys.* **25**, 127–165 (1994).
11. Coffin, J.M. HIV population dynamics in vivo: implications for genetic variation, pathogenesis and therapy. *Science* **267**, 483–489 (1995).
12. Frost, S.D.W. & McLean, A.R. Quasispecies dynamics and the emergence of drug resistance during zidovudine therapy of HIV infection. *AIDS* **8**, 323–332 (1994).
13. Haldane, J.B.S. A mathematical theory of natural and artificial selection. Part V. Selection and mutation. *Proc. Camb. Phil. Soc.* **28**, 838–844 (1937).
14. Crow, J.F. & Kimura, M. An introduction to population genetics theory. Anonymous (Harper and Row, New York, 1970).
15. Wright, S. Evolutionary and the genetics of populations. in *The Theory of Gene Frequencies* 512 (University of Chicago Press, Chicago, 1969).
16. Hudson, R.R. Gene genealogies and the coalescent process. in *Oxford Surveys in Evolutionary Biology* (ed. Futuyma, D. & Antonovics, J.) 1–44 (Oxford University Press, Oxford, 1990).
17. Tajima, F. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437–460 (1983).
18. Schuurman, 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. *J. Infect. Dis.* **171**, 1431–1437 (1995).
19. Richman, D.D. *et al.* HIV-1 mutants resistant to non-nucleoside inhibitors of reverse transcriptase arise in tissue culture. *Proc. Natl. Acad. Sci. USA* **88**, 11241–11245 (1991).
20. Havlir, D.V., Eastman, S., Gamst, A. & Richman, D.D. Nevirapine-resistant human immunodeficiency virus: Kinetics of replication and estimated prevalence in untreated patients. *J. Virol.* **70**, 7894–7899 (1996).
21. Richman, D.D. *et al.* Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy. *J. Virol.* **68**, 1660–1666 (1994).
22. Havlir, D.V. *et al.* Factors determining sustained antiviral response to nevirapine. 4th International HIV Drug Resistance Workshop. Sardinia, Italy, July 6–9, 1995 (Abstract).
23. Larder, B. Nucleosides and foscarnet — mechanisms. in *Antiviral Drug Resistance* (ed. Richman, D.D.) 169–190 (John Wiley & Sons, Chichester, 1996).
24. Richman, D.D., Grimes, J.M. & Lagakos, S.W. Effect of stage of disease and drug dose on zidovudine susceptibilities of isolates of human immunodeficiency virus. *J. AIDS* **3**, 743–746 (1990).
25. D'Aquila, R.T. Nucleosides and foscarnet — clincal aspects. in *Antiviral Drug Resistance* (ed. Richman, D.D.) 191–223 (John Wiley & Sons, Chichester, 1996).
26. Shirasaka, T. *et al.* Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy with azidothymidine, dideoxycytidine, and dideoxyinosine: An in vitro comparative study. *Proc. Natl. Acad. Sci. USA* **90**, 562–566 (1993).
27. Shafer, R.W. *et al.* Combination therapy with zidovudine and didanosine selects for drug-resistant human immunodeficiency virus type 1 strains with unique patterns of *pol* gene mutations. *J. Infect. Dis.* **169**, 722–729 (1994).
28. Shirasaka, T. *et al.* Emergence of human immunodeficiency virus type 1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides. *Proc. Natl. Acad. Sci. USA* **92**, 2398–2402 (1995).
29. Condra, J.H. *et al. In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374**, 569–571 (1995).
30. Gulick, R.M. *et al.* Potent and sustained antiretroviral activity of indinavir (IDV), zidovudine (ZDV) and lamivudine (3TC), XIth International Conference on AIDS, Vancouver, Canada, July 7–12, Abstr. Th. B 931 (1996) (Abstract).
31. Molla, A. *et al.* Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nature Med.* **2**, 760–766 (1996).
32. Kozal, M.J. *et al.* Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nature Med.* **2**, 753–759 (1996).
33. Lech, W.J. *et al. In vivo* sequence diversityof the protease of human immunodeficiency virus type 1: Presence of protease inhibitor-resistant variants in untreated subjects. *J. Virol.* **70**, 2038–2043 (1996).
34. Bozzette, S., McCutchan, J.A., Spector, S.A., Wright, B. & Richman, D.D. A cross-sectional comparison of persons with syncytium- and non-syncytium-inducing human immunodeficiency virus. *J. Infec. Dis.* **168**, 1374–1379 (1993).
35. Koot, M. *et al.* Prognostic value of HIV-1 syncytium-inducing phenotype for rate of $CD4^+$ cell depletion and progression to AIDS. *Ann. Intern. Med.* **118**, 681–688 (1993).
36. Richman, D.D. & Bozzette, S.A. The impact of syncytium-inducing phenotype of human immunodeficiency virus on disease progression. *J. Infect. Dis.* **169**, 968–974 (1994).
37. Muller, H.J. The relation of recombination to mutational advance. *Mutation Research* **1**, 2–9 (1964).
38. Kirchhoff, F. Greenough, T.C. Brettler, D.B., Sullivan, J.L. & Desrosiers, R.C. Absence of intact *nef* sequences in a long-term survivor with nonprogressive HIV-1 infection. *New Engl. J. Med.* **332**, 228–232 (1995).
39. Deacon, N.J. *et al.* Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* **270**, 988–991 (1995).
40. Leigh-Brown, A.J. Analysis of HIV-1 *env* gene sequences reveals evidence for a low effective number in the viral population. *Proc. Natl. Acad. Sci. USA* (in the press).
41. Delassus, S., Cheynier, R. & Wain-Hobson, S. Nonhomogeneous distribution of human immunodeficiency virus type 1 proviruses in the spleen. *J. Virol.* **66**, 5642–5645 (1992).
42. Kelly, J.K. An application of population genetic theory to synonymous gene sequence evolution in the human immunodeficiency virus (HIV). *Genet. Res.* **64**, 1–9 (1994).
43. Kelly, J.K. Replication rate and evolution in the human immunodeficiency virus. *J. Theor. Biol.* **180**, 359–364 (1996).

[1]*Centre for HIV Research, Institute of Cell, Animal and Population Biology University of Edinburgh, West Mains Road Edinburgh EH9 3JN, Scotland*

[2]*San Diego Veterans Affairs Medical Center and Departments of Pathology and Medicine, University of California San Diego, 9500 Gilman Drive La Jolla, California 92093-0679, USA*