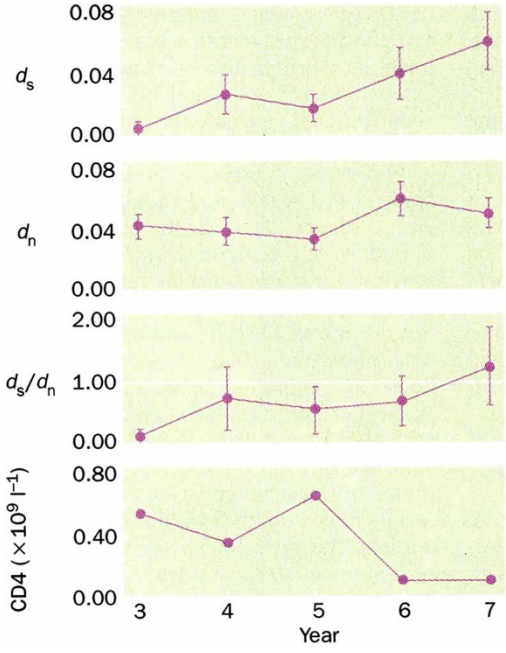
Causes of HIV diversity

SIR — The human immunodeficiency virus (HIV) is characterized by enormous genetic flexibility, which gives rise to drug resistance, escape from immune responses and failure of vaccination attempts. There is much discussion about the factors contributing to viral diversity in individual infections. It is clear that the high error rate of the reverse transcriptase[1] and the high turnover rate *in vivo*[2,3] generate vast numbers of different virus mutants. The diversity of viral quasispecies, however, is shaped by a combination of mutation and selection forces. The main selective forces that have been proposed to drive HIV diversity are the immune response, cell tropism and random activation of infected cells.

To quantify selection pressures on the HIV quasispecies, we have analysed the synonymous (amino-acid preserving) and nonsynonymous (amino-acid changing) nucleotide substitution pattern in the HIV-1 envelope gene of an infected haemophiliac patient followed since seroconversion for 7 years[4,5]. We compared 67 non-identical plasma-derived RNA sequences over a stretch of 231 bases, including the V3 loop and flanking regions. We derived the mean number of nucleotide substitutions per synonymous site, d_s, and per non-synonymous site, d_n, for all pairwise comparisons of sequences in the samples taken at years 3, 4, 5, 6 and 7 after infection. At year 0 the sequences were completely homogeneous.

The figure shows d_s, d_n, d_s/d_n and the CD4 cell count for years 3, 4, 5, 6 and 7 after seroconversion. As expected, d_s increases with time, due to accumulation of synonymous substitutions. The low value of d_s at year 3 may indicate that a selective sweep occurred in the viral population shortly before this time. It is interesting that d_n initially exceeds d_s, but thereafter increases much more slowly, if at all. At year 3 the d_s/d_n ratio is about 0.1, indicating strong positive selection for amino-acid change. As the infection progresses this selection pressure declines remarkably. By year 7 the ratio is about 1.25, indicating weak negative selection against amino-acid change. The pattern of increase in d_s/d_n agrees well with the pattern of decrease in the CD4 cell count. Whenever the CD4 cell count decreases, the d_s/d_n ratio increases.



Mean d_s, d_n and d_s/d_n for plasma-derived V3 sequences obtained from an infected haemophiliac patient and the corresponding CD4 cell count at years 3, 4, 5, 6 and 7 after seroconversion. The d_s/d_n ratios indicate initially strong selection pressure for amino-acid change, which then declines over time. The viral DNA sequences (GenBank accession numbers M84240–M84317) were extracted using nested polymerase chain reaction and direct sequencing[4]. The standard deviations of d_s and d_n are computed according to ref. 14. An approximation for the standard deviation of d_s/d_n is given by the propagated error $(\Delta d_s/d_n + \Delta d_n/d_n)d_s/d_n$.

A varying selection pressure is compatible with the notion that the immune responses select for viral diversity[6,7]. Early in the infection the immune system will respond strongly against common viral variants and hence favour rare mutants, thereby providing a strong positive selection pressure for diversification. As the immune system declines, this selection pressure should become weaker.

The pattern of decreasing positive selection, however, is also consistent with the notion that most viral diversity in the V3 region is caused by adaptation for various cell tropisms[8]. At seroconversion the patient carries a strongly homogeneous virus population (in the V3 region), then diversification occurs as HIV infects many different cell types and tissues in the body. Initially this would provide strong positive selection pressure which would decline when the virus has generated many variants with specific cell tropisms.

The observed pattern rejects the hypotheses that HIV diversity in V3 is purely neutral or simply generated by random activation of infected T-cell clones[9]. Such mechanisms would act equally on synonymous and non-synonymous substitutions and therefore predict a constant ratio of d_s/d_n during infection.

Phylogenetic analysis of the viral sequences reveals a division into two major lineages after year 3 (see ref. 4), which correspond to preferentially macrophage-tropic or T-cell-line-adapted phenotypes[10]. Comparison of the sequences in these two lineages (regardless to the sampling time point) reveals that the d_s/d_n ratio is significantly lower in the T-cell-line-adapted sequences than in the macrophage-tropic sequences (0.59 ± 0.37 and 1.39 ± 0.87, respectively). This result was independently confirmed by analysing 177 additional V3 sequences from 133 different donors[11]. For this dataset we found d_s/d_n ratios of 0.52 ± 0.16 for the T-cell-line-adapted sequences and 1.34 ± 0.45 for the macrophage-tropic sequences. This result can be explained if we assume that the immune response acts more strongly against the T-cell-line-adapted variant[12]. In the context of the cell-tropism hypothesis, we would have to assume that T-cell-line-adapted variants evolve a larger variety of cell tropisms, which seems unlikely. Therefore, it is plausible that immune selection has a greater effect on V3 diversity than cell tropism (but the two factors may be interlinked[13]).

This is the only patient studied so far in sufficient detail to provide estimates of d_s/d_n ratios (with reasonably small standard deviations) of individual points. More data of this kind are urgently needed if we are to understand the evolution of HIV in individual infections and its consequences for pathogenesis[6].

Sebastian Bonhoeffer
Edward C. Holmes
Martin A. Nowak
Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OXI 3PS, UK

1. Ji, J. & Loeb, L. A. *Virology* **185,** 323–330 (1994).
2. Wei, X. *et al. Nature* **373,** 117–122 (1995).
3. Ho, D. D. *et al. Nature* **373,** 123–126 (1995).
4. Holmes, E. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **89,** 4835–4839 (1992).
5. Zhang , L. Q. thesis, Univ. Edinburgh (1993).
6. Nowak, M. A. *et al. Science* **254,** 963–969 (1991).
7. Wolfs, T. F. *et al. Virology* **185,** 195–205 (199l).
8. Boyd, M. T. *et al. J. Virol.* **67,** 3649–3652 (1993).
9. Wain-Hobson, S. *Curr. Topics Microbiol. Immun.* **176,** 181–193 (1992).
10. Milich, L., Margolin, B. & Swanstrom, R. *J. Virol.* **67,** 5623–5634 (1993).
11. LaRosa, G. J. *et al. Science* **249,** 932–935 (1990).
12. Bou-Habib, D. *et al. J.Virol.* **68,** 6006–6013 (1994).
13. McKnight, A. *et al. J. Virol.* **69,** 3167–3170 (1995).
14. Nei, M. & Jin, L. *Molec. Biol. Evol.* **6,** 290–300 (1989).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.