Limitations of a Molecular Clock Applied to Considerations of the Origin of HIV-1

Bette Korber, James Theiler, Steven Wolinsky*

The evolutionary history of human immunodeficiency virus (HIV) is unfolding even as we study it, a consequence of the underlying genetic variation by mutation (*1*), recombination (*2, 3*), and frequent insertions and deletions (*4*). Studies of the evolutionary history of HIV enable us to extrapolate into the past and make estimates of the age of the epidemic, as well as make predictions of the potential for variation in the future that will affect vaccine development.

The HIV-1 main (M) group, although dominant in the global acquired immunodeficiency syndrome (AIDS) epidemic, is but one cluster in a complex array of simian and human lentiviruses (*5*). Other lentiviruses that can cause AIDS in humans include HIV-2, a distant cousin of HIV-1 that is common in West Africa and in India (*6*) and appears to have entered the human population through multiple zoonotic infections from simian immunodeficiency virus (SIV)–infected sooty mangabeys (*5*), and HIV-1 group O (*7*), a very distinctive "outlier" form of HIV-1 that is genetically more distant from the HIV-1 M group than a virus obtained from a chimpanzee caught in the wild in Gabon (CPZGAB) (*5, 8*). Nonetheless, it is the M group that has been preferentially amplified in the human population during the course of human events; whether this was due to chance or to biological differences in human lentiviruses remains unclear.

Differences can be found in greater than 25% of positions in envelope nucleotide sequences of diverse isolates within the M group. Genetic distance is represented in a phylogenetic tree as a branch length and is an estimate of how many mutational events actually occurred between two sequences. A simple tally of the number of changes between two sequences will underestimate the true genetic distance, because multiple mutational events may have occurred at any given site since divergence from a common ancestor. Phylogenetic methods use different ways to estimate the genetic distance between sequences and to organize a set of sequences into a hierarchy of ever more distantly related sequences (*9*). There are clearly defined associations among HIV-1 M group viruses that become apparent through phylogenetic analysis. This has led to the development of an alphabetical subtype taxonomy (*10, 11*) (Fig. 1), with designations A to J applied to the phylogenetically clustered lineages (or clades); subtypes G, I, and J are not yet fully characterized (*3*). The subtypes themselves can have complex patterns of subclusters (*12*), sometimes associated with the geographic origin of the virus. The relations of the subtypes within the M group of HIV-1 and their relative genetic distances are illustrated in a phylogenetic tree based on complete envelope sequences sampled from major global foci of the epidemic (Fig. 1).

Since HIV-1 sequences first began to accrue, researchers have been interested in estimating the age of the epidemic and the rate of viral evolution (*13, 14*) and in calibrating a molecular clock for HIV-1. Such a clock, if it were more uniform (*15*) than existential (*16*), would permit the use of contemporary sequences to examine hypotheses concerning the origins of HIV-1 in the human population. There are many potential problems with such back-projections, however (*17*). There clearly is a general trend toward greater evolutionary distances between viral sequences over time (Fig. 1). However, estimates of HIV-1 divergence rates are highly dependent on the region of the genome under study (*14*) (even different alignments of the same sequences can yield quite different divergence rates) and depend on the evolutionary model used to calculate the genetic distances (*9, 14*). A divergence rate can, in principle, be estimated by a linear regression analysis of genetic distances to an ancestral node, plotted as a function of the time of sampling. In practice, such plots provide rather dispersed clouds of points, with limited predictive power (*18*), even for an ideal molecular clock modeled as a Poisson process (*17*).

The calculation of divergence rates is further confounded by limitations inherent in phylogenetic reconstruction methods, which are ill-equipped to deal with evolutionary eccentricities that are well substantiated for HIV-1: namely, recombination and different evolutionary rates in different lineages. Phylogenetic analysis using recombinant fragments could yield branch lengths that do not reflect the evolutionary history of either parent strain. Recombination is probably contributing to HIV-1 evolution at every level, although it can be easily detected only when distinctive reference strains are available, such as two distinguishable strains infecting the same person (*19*), or two subtypes circulating in the same population (*2, 3*). Because recombination is common in populations harboring two or more subtypes, it is reasonable to assume that recombination is also a factor in the generation of HIV-1 diversity in a population with a preponderance of a single subtype, despite the fact that it cannot be readily detected. There is also suggestive evidence that recombination events occurred before the formation of the subtypes (*20, 21*), and such events would obscure deeper ancestral associations.

Whereas recombination muddles the branching patterns of phylogenetic reconstructions, different rates of evolution in sublineages disrupt the relation between evolutionary distances and time. In different individuals, differences in rates of evolution may be influenced by differences in host-mediated selection pressures. Early in infection, a relatively narrow range of genetic variants is observed, which subsequently diversifies genetically, biologically, and antigenically (*22, 23*). The high mutation frequency of HIV-1 (*1*) coupled with a continuous high rate of virus production (*24*) yields an extensive reservoir of variants within a single host, providing a fertile ground for natural selection (*25*), the critically important means by which variants that escape immunological detection (*26*) and drug-resistant strains (*27*) arise. The viral quasispecies can also expand into new cellular populations by acquiring mutations that facilitate adaptation (*28*). During the persistent stage of infection, it is not unusual to find HIV-1 envelope sequences from a single individual that differ in more than 10% of their nucleotide positions. Distinctly slow rates of viral evolution have been noted in individuals who rapidly progress to disease. Individuals who progress normally or slowly tend to have a substantial increase in viral diversity with time, even during the first few years of their infection. In contrast, individuals who get AIDS within a few years of their initial infection generally have a more homogeneous viral population until their death (*22*). There is also evidence for different rates of evolution in different subtypes, at least in the third vari-

B. Korber is in the Theoretical Division (T-10), Los Alamos National Laboratory, Los Alamos, NM 87545, USA, and is a visiting faculty member at the Santa Fe Institute, Santa Fe, NM 87501, USA. J. Theiler is in the Space and Remote Sensing Sciences Group (NIS-2), Los Alamos National Laboratory, Los Alamos, NM 87545, USA. S. Wolinsky is in the Department of Medicine, Northwestern University Medical School, Chicago, IL 60611, USA.

*To whom correspondence should be addressed.

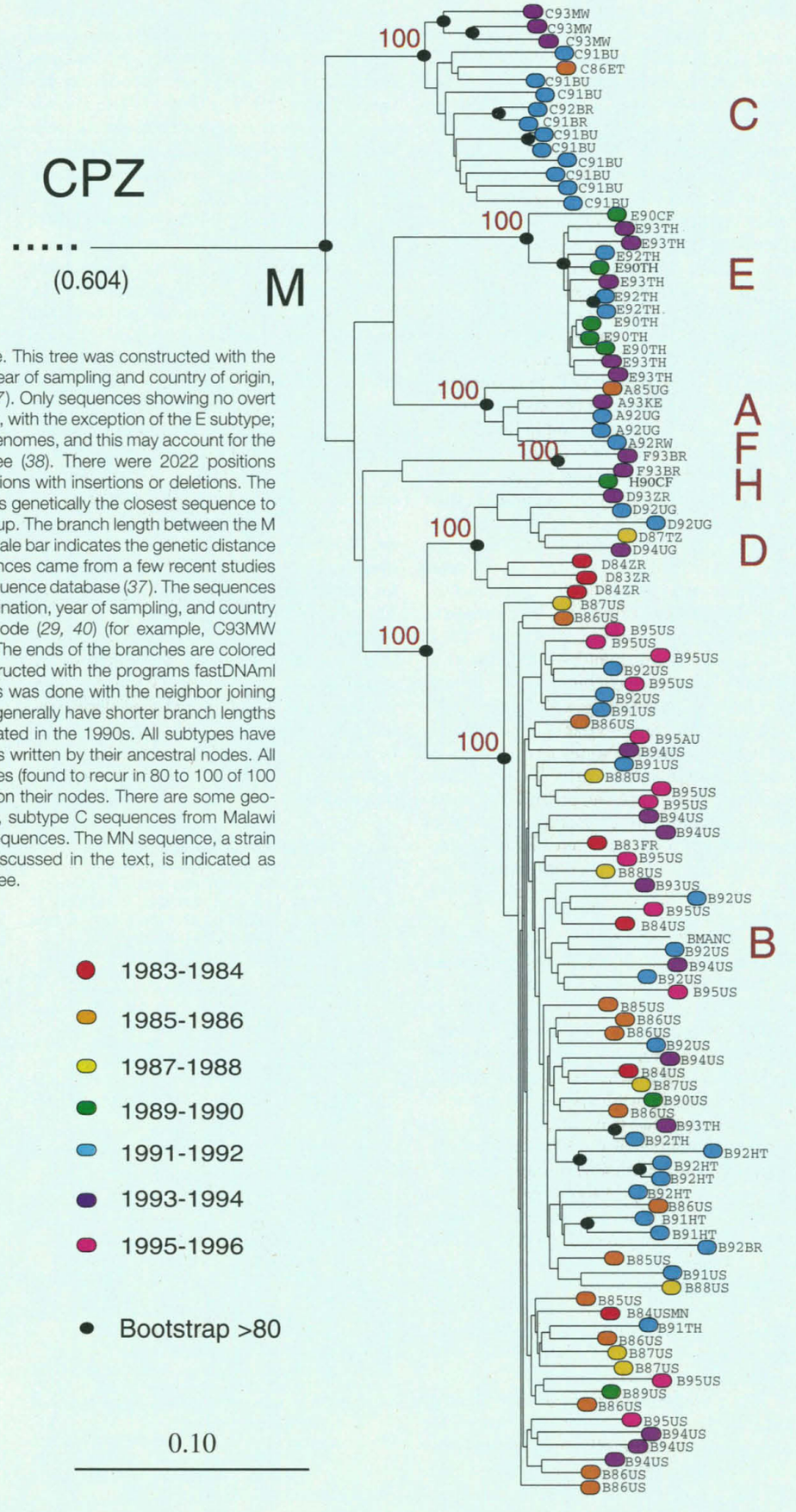

Fig. 1. An HIV-1 envelope maximum likelihood tree. This tree was constructed with the 113 complete envelope sequences with a known year of sampling and country of origin, which were readily available in the HIV database (*37*). Only sequences showing no overt evidence of interclade recombination were included, with the exception of the E subtype; all known E subtype sequences have E-A mosaic genomes, and this may account for the association of E and A subtypes seen in the tree (*38*). There were 2022 positions remaining in the alignment after elimination of positions with insertions or deletions. The CPZGAB sequence, a chimpanzee sequence that is genetically the closest sequence to HIV-1 M group sequences, was used as the outgroup. The branch length between the M group node and the CPZ sequence is 0.604. The scale bar indicates the genetic distance of the branch lengths. Many of the included sequences came from a few recent studies (*36*, *39*), and basic alignment came from the HIV sequence database (*37*). The sequences in the tree, as shown here, retain their subtype designation, year of sampling, and country of origin, as reflected by the two-letter country code (*29, 40*) (for example, C93MW indicates C subtype, sampled in 1993, in Malawi). The ends of the branches are colored to reflect the time of sampling. The tree was constructed with the programs fastDNAml (*41*) and DNArates (*42*, *43*), and bootstrap analysis was done with the neighbor joining package of Phylip (*44*). Taxa isolated in the 1980s generally have shorter branch lengths from a shared ancestral node than do viruses isolated in the 1990s. All subtypes have bootstrap values of 100/100, indicated by the 100s written by their ancestral nodes. All very reliable branch points with high bootstrap values (found to recur in 80 to 100 of 100 boostrap replicates) are indicated by black marks on their nodes. There are some geographic associations within subtypes; for example, subtype C sequences from Malawi form a subcluster, as do some Haitian subtype B sequences. The MN sequence, a strain that has been used for vaccine studies and is discussed in the text, is indicated as B84USMN, the 14th taxa from the bottom of the tree.

able domain (V3) of the envelope protein, the most heavily sequenced region of the viral genome (*29*) and an important functional and immunogenic region. In the D subtype, the V3 loop is mutating relatively rapidly, in the A and B subtypes it is mutating at a moderate pace, and in the C subtype it is changing slowly (*29, 30*).

Given these complicating factors, attempts to discern when the virus was introduced into the human population can be treacherous. Nonetheless, estimating the age of a common ancestor of the AIDS pandemic is of great interest. In this regard, the sequence of a sample from the Democratic Republic of the Congo (formerly Zaire) obtained in 1959 (ZR.59) is informative (*20*). Phylogenetic analyses established the authenticity of the sample and placed the origin of the ZR.59 sequence very near the ancestral node of the B, D, and F clades of the M group, anchoring the node in time (*21*). It suggests that the divergence found in these clades arose after 1959; the extent of the diversification over the last 40 years is striking and provides a disturbing indication of what the future holds. The extensive within-subtype variation of at least the B and D clades has accumulated in a matter of decades; by extension, the virus may be able to rapidly acquire new levels of diversity by superimposing new variants onto the already diverse array of forms now well established throughout the world. Thus, vaccines we devise must have the potential to cross-protect against an extraordinary array of variants, and we should consider using vaccine strategies that would optimize the potential for cross-strain protection, both within and between clades (*31*).

The analyses of the 1959 Congo sequence also suggest that the precursor of the modern global epidemic and dispersion of the M group viruses in the human population occurred not long before 1959, probably within decades (*20*). This is only a rough estimate and does not date the nonhuman primate lentivirus precursor of HIV-1, an event or series of events that may have occurred hundreds or thousands of years earlier. A divergence rate can be calculated from Fig. 1 (*18, 32*) and used as an alternative strategy to estimate the timing of ancestral branch points. Given the caveats discussed above, such estimates are tentative; however, if one carries through the analysis to estimate the year the B and D clades diverged and the M group arose, the results are in accord with the conclusions of the study of the 1959 Congo sequence (*20, 32*).

Better understanding of the time frame of the HIV-1 epidemic is only one reason for attempting to define the phylogenetic relations and the potential for variation within the HIV-1 genome. To guide vaccine strategies it is critically important to anticipate the evolution of the virus and to understand genetic patterns in contemporary variants, particularly the geographic association between clades (*10, 11*) (Fig. 1). The B clade viruses of North America have genetic relations that approximate a star phylogeny, where the viral sequences radiate outward from an ancestral node in a phylogenetic tree; thus, no one virus is truly representative of the epidemic (Fig. 1). However, there is a line of thinking that persists among some AIDS researchers that a sequence in the V3 domain of the HIV-1 vaccine-candidate strain MN, Ile-Gly-Pro-Gly-Arg-Ala-Phe [an antigenic domain that is a good neutralizing antibody target for laboratory-adapted strains, but a far less potent target in primary viruses (*33*)], is the representative form of contemporary viruses circulating in North America (*34*). Because the HIV-1 V3 domain sequences are diversifying, the proportion of viruses that closely match the MN V3 motif are inexorably diminishing (*29, 35, 36*). The overall extent of viral variation, as illustrated for the full-length envelope sequences (Fig. 1), is likely to be problematic for any vaccine candidate. But unless effective counterstrategies can be devised, the amplification and diversification of the HIV-1 M group viruses in the human population during the last half-century is likely to be transcended by relentless amplification and diversification into the next; a sobering perspective provided by a look back into the future.

REFERENCES AND NOTES

1. J. M. Coffin, *Science* **267**, 483 (1995); B. D. Preston, B. J. Poiesz, L. A. Loeb, *ibid.* **242**, 1168 (1988); K. Bebenek, J. Abbotts, S. H. Wilson, T. A. Kunkel, *J. Biol. Chem.* **268**, 10324 (1993).
2. E. C. Sabino *et al.*, *J. Virol.* **68**, 6340 (1994); D. L. Robertson, B. H. Hahn, P. M. Sharp, *Mol. Evol.* **40**, 249 (1995); F. E. McCutchan, M. O. Salminen, J. K. Carr, D. S. Burke, *AIDS* **10**, S13 (1996); M. Cornelissen *et al.*, *J. Virol.* **70**, 8209 (1996).
3. D. L. Robertson, F. Gao, B. H. Hahn, P. M. Sharp, in *Human Retroviruses and AIDS*, B. Korber *et al.*, Eds. (Los Alamos National Laboratory, Los Alamos, NM, 1997), part III, pp. 25–30, http://hiv-web.lanl.gov/HTML/reviews/hahn.html.
4. S. M. Wolinsky *et al.*, *Science* **255**, 1134 (1992); C. Wilis, G. Myers, A. Farmer, *AIDS Res. Hum. Retroviruses* **12**, 1383 (1996).
5. P. M. Sharp, D. L. Robertson, F. Gao, B. H. Hahn, *AIDS* **8**, S27 (1994).
6. F. Clavel *et al.*, *Nature* **324**, 691 (1986); M. Grez *et al.*, *J. Virol.* **68**, 2161 (1994); B. Korber, E. Allen, A. Farmer, G. Myers, *AIDS* **9** (suppl. A), S5 (1995); P. J. Kanki, M. Peeters, A. Gueye-Ndiaye, *ibid.* **11** (suppl. B), S33 (1997).
7. P. Charneau *et al.*, *Virology* **205**, 247 (1994); I. Lousert-Ajaka *et al.*, *J. Virol.* **69**, 5640 (1995).
8. T. Huet *et al.*, *Nature* **345**, 356 (1990).
9. D. L. Swofford, G. J. Olsen, P. J. Waddell, D. M. Hillis, in *Molecular Systematics*, D. M. Hillis, C. Moritz, B. K. Mable, Eds. (Sinauer, Sunderland, MA, ed. 2, 1996), pp. 407–514.
10. J. Louwagie *et al.*, *AIDS* **7**, 769 (1993); Workshop Report from the European Commission (DG XII, INCO-DC) and the Joint United Nations Programme on HIV/AIDS, *AIDS* **11**, 17 (1997).
11. T. Leitner, B. Korber, D. Robertson, F. Gao, B. Hahn, in (*3*), pp. 19–26, http://hiv-web.lanl.gov/HTML/reviews/Thomas.html.
12. F. E. McCutchan *et al.*, in *The Proceedings of the Fifth Annual HIV Dynamics and Evolution Meeting*, Santa Fe, NM, 18 to 20 April 1998, B. Korber, Ed. (The Santa Fe Institute, Santa Fe, NM, 1998), vol. 5, p. 23, www.santafe.edu/hiv5/mccutchan.html; M. Peeters *et al.*, *ibid.*, p. 73, www.santafe.edu/hiv5/peeters.html.
13. W.-H. Li, M. Tanimura, P. M. Sharp, *Mol. Biol. Evol.* **5**, 313 (1988); P. Sharp and W.-H. Li, *Nature* **333**, 315 (1988); T. Gojobori, E. N. Moriyama, M. Kimura. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 10015 (1990).
14. T. Leitner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10864 (1996).
15. "Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external." Isaac Newton, *Principia* (1687), F. Cajori, transl. (Univ. of California Press, Berkeley, CA, 1947), p. 6; "Wrong again, Isaac," paraphrased by P. Ginsparg (personal communication), from Albert Einstein, *Special Theory of Relativity* (1905).
16. "One hand only. Finest of fine black darts. It advances by fits and starts. No tick. Leaps from dot to dot with so lightning a leap that but for its new position it had not stirred. Whole nights may pass as but a fraction of a second or any intermediate lapse of time soever before it flings itself from one degree to the next." Samuel Beckett, author's translation, *Ill Seen, Ill Said* (Grove, New York, 1981), p. 76.
17. D. M. Hillis, B. K. Mable, C. Moritz, in (*9*), pp. 515–543.
18. The use of the distance to a common M group ancestral node to estimate the year of sampling of a sequence gives a root mean square error of 8.7 years. Given that the virus that causes AIDS was identified only 15 years ago, this is not much predictive power. This calculation was specifically based on a linear regression analysis of genetic distance between contemporary sequences and the estimated ancestral origin of the HIV M group, shown in Fig. 1, plotted as a function of the year of sampling. The linear correlation coefficient (Pearson's r) is 0.282 for the full set of data in Fig. 1, with a Student's t probability of 0.002, that is, a highly significant but weak correlation. Using a regression analysis to estimate the rate of divergence and a bootstrap approach to estimate the error (*45*), we calculated annual divergence rates of 0.00169 per year (95% confidence interval of 0.0011 to 0.0022). These rates are lower than numbers given in the past (approaching 1% per year) [C. Kuiken and B. Korber, *AIDS* **8**, S73 (1994)], perhaps because these calculations were based on sequences of the full-length envelope protein gp160, including the conserved regions, and typically shorter variable domains of envelope have been studied.
19. M. O. Salminen *et al.*, *J. Virol.* **71**, 2647 (1997); R. S. Diaz *et al.*, *ibid.* **69**, 3273 (1995).
20. T. Zhu *et al.*, *Nature* **391**, 594 (1998).
21. The branch location of the ZR.59 sample, between the B, D, and F clades, could have been the result of ZR.59 being a mosaic genome; however, the branch length was extremely short between the ancestral node of these clades and the ZR.59 sequence, indicating that it was simply close to the ancestral sequence. There was no evidence that the ZR.59 sequence was a mosaic when it was analyzed carefully for that possibility with a sliding window approach. However, evidence for a preclade recombination event emerged from the study. The B and D subtypes are generally the most closely associated subtypes throughout the genome. But in the fragments of the genome where sample ZR.59 was sequenced, B and F subtype sequences were most closely associated. [Additional evidence for early recombination events is that the subtypes cluster with different patterns of association in phylogenetic trees on the basis of different regions of the genome (*11*)].
22. S. Wolinsky *et al.*, *Science* **272**, 537 (1996); S. Ganeshan, R. Dickover, B. Korber, Y. Bryson, S. Wolinsky, *J. Virol.* **71**, 663 (1997); S.-L. Liu *et al.*, *ibid.*, p. 4284; E. Delwart *et al.*, *ibid.*, p. 7498; V. V. Lukashov *et al.*, *ibid.* **69**, 6911 (1995).
23. M. A. Nowak *et al.*, *Nature* **375**, 606 (1995).

24. A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, D. D. Ho, *Science* **271**, 1582 (1996).
25. M. Eigen, *Trends Microbiol.* **4**, 216 (1996).
26. P. Borrow *et al.*, *Nature Med.* **3**, 205 (1997); D. A. Price *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1890 (1997).
27. H. F. Gunthard *et al.*, *J. Virol.* **72**, 2422 (1998); J. Hammond, B. A. Larder, R. F. Schinazi, J. W. Mellors, in (*3*), pp. 206–249.
28. R. Doms and J. P. Moore, in (*3*), p. 1, http://hiv-web.lanl.gov/HTML/reviews/doms.html.
29. P. Dighe, B. Korber, B. Foley, in (*3*), p. 74, http://hiv-web.lanl.gov/HTML/97compendium.html.
30. B. T. M. Korber, K. MacInnes, R. Smith, G. Myers, *J. Virol.* **68**, 6730 (1994); C. L. Kuiken, B. T. Foley, E. Guzman, B. T. M. Korber, in *Molecular Evolution of HIV*, K. Crandall, Ed. (Johns Hopkins Univ. Press, Baltimore, MD, in press).
31. H. Cao *et al.*, *J. Virol.* **71**, 8615 (1997); G. Ferrari, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1396 (1997).
32. Using a method that determines the best fit line for given data points including error estimates in both dimensions, one can determine how different assumptions of error influence the slope of a line [W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, ed. 2, 1992), section 15.3, p. 666]. Use of the Poisson error as the fixed error for the distance and variation of the time error yields a range of quite different slopes (and as a consequence, the estimated year of origin of the M group) for the data shown in Fig. 1. This ranges from the extreme estimates of 1917 (no error in the time axis) to 1972 (at the limit of large error in the time axis, the equivalent to a fit with no error in the distance axis). If one assumes that the year of sampling could be confounded by a small pool of latently infected cells, carrying virus that originated some years before sampling, the predicted year of origin is intermediate. For example, if one assumes a 6-year error in the time axis, the estimated year of an ancestral sequence of the M group is 1942 (*44*).
33. J. P. Moore and D. D. Ho, *AIDS* **9** (suppl. A), S117 (1995); M. P. D'Souza, D. Livnat, J. A. Bradac, S. H. Bridges, *Infect. Dis.* **175**, 1056 (1997); S. Beddows *et al.*, *J. Gen. Virol.* **79**, 77 (1998).
34. P. Berman *et al.*, *J. Virol.* **66**, 4464 (1992); P. Berman *et al.*, *J. Infect. Dis.* **176**, 384 (1997); "FDA authorizes first full testing for HIV vaccine," *New York Times*, 4 June 1998, p. A1.
35. B. Korber *et al.*, *AIDS Res. Hum. Retroviruses* **8**, 1461 (1992); S. D. Rencher, T. D. Lockey, K. S. Slobod, J. L. Hurwitz, *ibid.* **13**, 527 (1997).
36. R. Connor *et al.*, *ibid.* **72**, 1552 (1998).
37. Los Alamos, NM (1997). Theoretical Biology and Biophysics, Los Alamos National Laboratory, http://hiv-web.lanl.gov/HTML/ENV-index.html.
38. J. K. Carr *et al.*, *J. Virol.* **70**, 5935 (1996); F. Gao *et al.*, *ibid.*, p. 7013.
39. F. Gao *et al.*, *ibid.* **70**, 1651 (1996); F. McCutchan *et al.*, *AIDS Res. Hum. Retroviruses* **14**, 329 (1998).
40. Two-letter country codes indicating the origins of the viruses included in the envelope tree in Fig. 1 are as follows: Malawi (MW), Burundi (BU), Ethiopia (ET), Brazil (BR), Central African Republic (CF), Thailand (TH), Uganda (UG), Kenya (KE), Rwanda (RW), Zaire (ZR), United States (US), Australia (AU), France (FR), and Haiti (HT).
41. G. J. Olsen *et al.*, *Comput. Appl. Biosci.* **10**, 41 (1994).
42. FastDNAml and DNArates were written by Gary Olsen and colleagues at the Ribosomal Database Project (RDP) at the University of Illinois at Urbana-Champaign, available by anonymous ftp from ftp://rdp.life.uiuc.edu/.
43. To create this tree, we generated a preliminary maximum likelihood tree which was then used as input for the DNArates program (*41*) that produced the maximum likelihood estimate of the rate of nucleotide substitution at every position in the alignment, allowing nine categories of rates. The preliminary tree and rates and categories were used to create a new tree, allowing global rearrangements and adjusting of branch lengths. Maximum likelihood was used to generate the tree to get the most accurate branch lengths possible (*9, 11*). It took 3 weeks of computer time on a SPARC workstation to generate this tree, so bootstrap analysis was done by using the neighbor joining package of Phylip with the maximum likelihood model to calculate the genetic distances (*43*).
44. J. Felsenstein, *Evolution* **39**, 783 (1985); *Cladistics* **5**, 164 (1989).
45. B. Korber and J. Theiler, unpublished observations.
46. We thank C. Kuiken, T. Leitner, K. Rock, and B. Foley for helpful discussions, R. Thakallapally and A. Ploss for their help with the preparation of this manuscript, and P. Ginsparg for suggestions concerning the text. B.K. was supported by interagency agreements between the U.S. Department of Energy (DOE) and NIH, 3-Y01-AI-70001-13 and Y1-A1-4058-03; J.T. by the DOE; and S.M.W. by grants RO1-HD-31756 and PO1-AI-42454 from NIH, and a gift from an anonymous foundation.

HIV Treatment Failure: Testing for HIV Resistance in Clinical Practice

Luc Perrin and Amalio Telenti

In a recent commentary on AIDS therapy, the phrase "Failure isn't what it used to be . . . but neither is success" was coined (*1*). By now we should be used to issues concerning the human immunodeficiency virus (HIV) being always more complex than expected. Understanding of HIV pathogenesis indicated that early pharmacological intervention would give the best chance at preserving the integrity of the immune system and possibly eradicating the virus. These concepts provided the impetus to treat a large proportion of HIV-infected individuals with a combination of antiretroviral drugs [highly active antiretroviral therapy (HAART)], resulting in a dramatic reduction in AIDS-related morbidity and mortality (*2*). However, viral eradication is not achievable with current strategies (*3*), and the shift in treatment paradigm to one of long-term viral suppression has led to the challenge of ensuring continuous treatment benefit and avoiding failure (*4*).

Failure has generally been defined in virological terms—the inability to achieve complete suppression of viral replication. The factors leading to this type of failure are straightforward: poor adherence to HAART, prior exposure to antiretroviral drugs in mono- or bi-therapy, the sequential addition of drugs to a failing regimen, and counteractive interactions among the drugs used (*5*)—nothing new for those who witnessed the early days of antituberculosis chemotherapy in the 1950s and 1960s. However, treatment failure is not only viral resistance.

In fact, definition of failure or success of treatment is a far more complex phenomenon (Fig. 1). In real life, there are individuals who experience an optimal response to treatment, as shown by effective viral suppression and ensuing immune recovery (6) (Fig. 1A), but there are others with increasing CD4 cell counts in the presence of ongoing viral replication (*7*) (Fig. 1B), or blunted immune recovery despite viral control (Fig. 1C), and finally complete treatment failure (Fig. 1D). Analysis of the Swiss HIV cohort study database of HIV-1–infected individuals on HAART indicates that an estimated 40% of the participants present the constellation described in Fig. 1A, 40% in Fig. 1B, 5% in Fig. 1C, and 15% in Fig. 1D. We need to understand better what each situation represents clinically and what each implies for the current models of HIV immunopathogenesis (*8*). Finally, we have to learn more about the mechanisms by which current antiretroviral drugs exert their remarkable effect on HIV disease despite widespread drug resistance (*7, 9*). In particular, the frequent observation of increasing CD4 cell counts in individuals maintaining high viremia levels needs to be explained, because it may yield clues regarding issues such as viral fitness, resetting in the steady state of CD4 cell turnover, and the possible action of protease inhibitors on nonviral targets participating in the mechanisms of CD4 T cell depletion.

Resistance is a widespread problem. Although treatment failure is a complex phenomenon, viral resistance indeed remains a major issue. It affects up to 30 to 50% of all individuals under HAART (*7, 9*) and also might be transmitted (*10*). Once multidrug resistance is present, regaining control of viremia becomes difficult because no effective "salvage" strategy has been devised.

However, testing of HIV resistance is not straightforward because the best analysis strategies have not been defined and remain the topic of intense clinical investigation. Central to its complexity are the phenomenon of HIV quasi-species (the simultaneous presence in a patient of a swarm of viral variants), the extent of cross-resistance among antiviral drugs, the existence in each individual of archival HIV DNA

The authors are at the Laboratory of Virology, Division of Infectious Diseases, Geneva University Hospital, 1211 Geneva 14, and Division of Infectious Diseases, Centre Hospitalier Universitaire Vaudois, 1011 Lausanne, Switzerland.