Points of View

Syst. Biol. 44(1):77–92, 1995

The AIDS Pandemic Is New, But Is HIV New?

DAVID P. MINDELL,[1,4] JEFFREY W. SHULTZ,[2] AND PAUL W. EWALD[3]

[1]*Department of Biology and Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109-1079, USA*
[2]*Department of Entomology, University of Maryland, College Park, Maryland 20742, USA*
[3]*Department of Biology, Amherst College, Amherst, Massachusetts 01002, USA*

The AIDS pandemic is a new problem for humans, but it is unclear whether the human immunodeficiency virus (HIV) giving rise to AIDS is also new to humans. Either HIV has recently infected humans, in which case we have a new virus and a new disease, or HIV infected humans long ago (being mild and/or restricted in range until recently), in which case we have an old virus and a new disease. There are precedents for each scenario among known viruses causing diseases (see Shope and Evans, 1993). The new virus and old virus scenarios have profoundly different implications for understanding the mechanisms of HIV propagation and the etiology of AIDS, for combating AIDS, and potentially for efforts to prevent future epidemics. The terms "new" and "old" are ambiguous beyond denoting relative age; for purposes of this article we consider a new virus one that has infected its host species within the past 50 years or so.

The first view, that HIV has only recently contacted humans, entails recent cross-species transmission of a simian immunodeficiency virus (SIV) from one or more nonhuman primates and represents the current conventional wisdom (Dietrich et al., 1989; Doolittle, 1989; Allan et al., 1991; Fox, 1992; Myers et al., 1992, 1993b; Hirsch et al., 1993; Temin, 1993; Myers and Korber, 1994). However, some have suggested that certain rural African populations of humans may have been infected with an immunodeficiency virus for many decades, centuries, or even millennia (Montagnier, 1985; Hahn, 1990; McClure, 1990), and Ewald (1991, 1994) described an evolutionary model in which virulent strains are placed at a selective advantage by higher rates of sexual partner change.

Understanding HIV origins is of general interest to systematists. Viruses evolve by descent with modification like any other group of organisms, and systematists will become increasingly involved in attempts to understand their complex histories as more of their DNA sequences become available. Systematists working on viruses need to consider distinctive features of viral evolution, including extremely high rates of molecular sequence evolution, subsequent high levels of within-population sequence variability (variously described as yielding species swarms or quasi species), evolutionary rates that vary depending on the species of host and type of cell infected, potential for

[4] E-mail: mindell@umich.edu.

recombination when representatives of different viral lineages infect the same host cell, and potential biases in the sampling of host species.

Our objectives here are to (1) assess the evidence used in support of the "new virus" hypothesis, (2) present our own phylogenetic analyses of representative viral taxa, (3) estimate the most-parsimonious evolution of the character "virus host" in the study taxa, and (4) comment on methodological issues in the systematics of viruses.

Although it is not possible to reject either hypothesis, we conclude that any consensus favoring the "new virus" hypothesis is not justified on the basis of current evidence and that the "old virus" hypothesis remains a viable alternative.

Phylogenetic Analyses of HIV Origins

Previous Studies

HIVs and SIVs are retroviruses, a group characterized by the ability to reverse transcribe RNA into DNA. HIV and SIV genomes are about 10 kilobases in size and contain at least nine recognizable genes. Previous phylogenetic analyses of primary HIV and SIV lineages have used, variously, the *pol* and *gag* genes. The *pol* gene encodes reverse transcriptase and endonuclease enzymes, whereas *gag* encodes capsid protein (*gag* p24), which forms a shell around the viral RNA, and several internal proteins (*gag* p15, *gag* p18) functioning in viral reproduction (Stine, 1993; Hahn, 1994). The regions of *pol* encoding reverse transcriptase and endonuclease comprise the most slowly evolving regions of the genome (McClure et al., 1988). Previous studies have found that HIVs and SIVs form a monophyletic group, with immunodeficiency viruses from the domestic cat (FIV), sheep (VISNA), and horse (EIAV) being closely related to the primate immunodeficiency viruses (Doolittle et al., 1989; Yokoyama, 1991).

Because HIVs are parasites, the question of their origins includes two issues: (1) the phylogeny of the viruses and (2) the history of virus transmission among host species. Parsimony may be used in addressing both of these questions. Parsimony is used in estimating the history of host shifts for viral taxa by minimizing the number of ad hoc assumptions of host shift on a phylogenetic tree. This minimization of ad hoc assumptions of host shift is justified on the same logical basis as is the minimization of ad hoc assumptions of character convergence (homoplasy) in phylogenetic analyses using a parsimony criterion (see Farris, 1983). Further, most viruses are narrowly host specific and are unable to survive immune system surveillance in new host species. These characteristics are also consistent with a parsimony approach to assessing virus transmission history.

However, patterns of phylogeny for extant viruses and patterns of their history of transmission need not be congruent. For example, the appearance of a virus from a particular host species as basal in a phylogeny could be the result of extinction of lineages from the true early host, which "gave" the virus to the current host recently, or could result from a lack of sampling from the true early host species. Also, phylogenetic trees for viral taxa do not indicate whether a particular host species was a virus donor or a virus recipient. The trees simply denote hypothesized lineage-splitting events among viruses. For these reasons, phylogenetic trees for viruses must be interpreted with caution in assessing the history of virus host shifts.

Use of phylogenetic analyses to support the "new virus" hypothesis is common in the literature. For example, Myers et al. (1993b:126) presented a phylogenetic hypothesis for HIVs and SIVs based on a 648-nucleotide section of the *gag* p24 gene (Fig. 1) and said that this "phylogenetic tree analysis . . . strongly supports the hypothesis of the simian origin of AIDS" (the "new virus" hypothesis). However, we see no such support in the tree itself. There are both HIVs and SIVs on either side of the first bifurcation event within the tree, and various HIVs could have either descended from or given rise to various SIVs. For example, sooty mangabey viruses (ancestral to SIVsmmh4 and SIVsmpbj in Fig. 1) might have given rise to HIV2s (HIV2rod, HIV2d205), or con-

versely an ancestral HIV2 lineage may have given rise to the sooty mangabey viruses. The tree itself is consistent with either scenario. Phylogenetic trees show purported sister relationships among extant lineages but do not denote ancestor–descendant relationships among those extant lineages.

Rather, support for the "new virus" hypothesis is nonphylogenetic and circumstantial, being rooted in unsupported assumptions that (1) surveys of HIV presence in human blood samples collected before 1980 are based on reliable and sufficiently large samples, (2) high virulence denotes recency of the infection of the host species, (3) all HIVs are virulent, giving rise to AIDS, and (4) SIVs are not virulent in their natural hosts. Regarding the first point, researchers have pointed out that HIV seropositivity assays for human blood samples collected prior to the 1980s are largely negative, and assays for blood samples collected during the 1980s in particular show increasing levels of positivity (Grmek, 1990; see Myers et al., 1993b, and references therein). These facts are consistent with the notion of the first human infection occurring during the middle portion of this century, but they do not refute the alternative hypothesis that HIVs were present in one or more small and possibly isolated human populations not represented in pre-1980 or even pre-1959 blood samples (1959 is the collection date for the earliest known seropositive sample). As dramatic as the seropositivity surveys are, their obvious geographic and quantitative sampling limitations compromise their ability to delineate the timing of infection. In a review of retrospective seropositivity surveys, Myers et al. (1993b) discussed the data available, which comes from just seven geographic locations across Africa involving hundreds or thousands of human blood samples. Not surprisingly, vast regions of Africa with millions of human inhabitants are not represented in retrospective seropositivity surveys. HIVs are lentiviruses, which are characterized as a group by their potential for long periods of latency with no visible effects on hosts. HIVs present, though perhaps inconspicuous, in small isolated human populations or demes could readily have been missed by limited pre-1959 (or even pre-1980) blood samplings. Current high levels of virulence, particularly for HIV1s, may have been generated by recent changes in host behavior relating to virus transmission opportunities (Ewald, 1994).

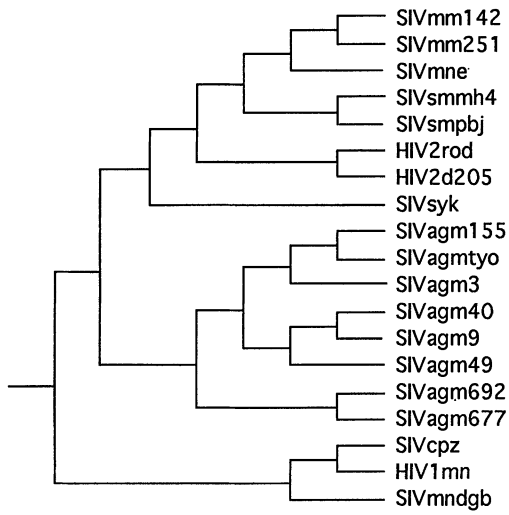

FIGURE 1. Phylogenetic tree topology from Myers et al. (1993b) based on immunodeficiency viral *gag* p24 sequences for 648 nucleotide positions. Tree is rooted at the midpoint of the greatest patristic distance between two terminals. Host species and abbreviations for viruses are defined in the Appendix.

Regarding the other three assumptions, Doolittle (1989:339) noted the circumstantial view in pointing out that "primary hosts all seem to be healthy," whereas the likely secondary hosts are not healthy. Myers et al. (1992:373) stated

> given the pervasiveness of SIVs in diverse African monkey populations, and the relative newness of HIV in human populations, the hypothesis of a recent simian origin of human AIDS through one or more events of cross-species transmission has gained widespread acceptance over the past few years.

However, if the "relative newness of HIV in human populations" is "given," how can one reach any other conclusion? Often, authors simply note the sister relationship be-

tween viruses isolated from different host species and presuppose the direction of infection to be from nonhumans to humans. These assumptions should not be accepted uncritically. Other researchers have been more cautious in inferring direction of infection from phylogenetic analyses. For example, Hirsch et al. (1989:391) noted that their

> data cannot exclude the possibility that HIV2 from a human was passed to a sooty mangabey and subsequently evolved as SIVsm. . . . [S]equences of older HIV2 and SIV isolates (from mangabeys or other species) are required to resolve these issues.

Although direction of infection cannot be inferred from a sister relationship between two viruses from different host species, accurate tree topology is crucial for estimating the number of host shifts that have occurred and the most-parsimonious sequence of host shift events. Many aspects of HIV and SIV relationships are poorly resolved, particularly the earliest divergences involving five lineages: (1) HIV1s/SIVcpz, (2) HIV2s/SIVsm/SIVmac, (3) SIVmnd, (4) SIVagms, and (5) SIVsyk. HIVs, as currently named, are clearly not monophyletic. The two primary HIV types, HIV1 and HIV2, each include representative strains (or taxa) that are more closely related to one or more SIVs than they are to other HIVs. As a corollary, SIVs are also not monophyletic. Whether HIV1s are monophyletic and HIV2s are monophyletic is less clear. SIVcpz, previously placed as sister to all HIV1s (Huet et al., 1990), may belong inside an HIV1 clade when divergent HIV1s are included in the analyses. Similarly, SIVs from sooty mangabeys and macaques are often, though not always, placed within the HIV2 clade.

Clearly, there have been multiple host species shifts by the viruses, and given a nonsister relationship between the HIV1 and HIV2 types, proponents of the "new virus" hypothesis must invoke at least two recent and independent human infections with quite different viruses, each capable of causing AIDS. HIV1s and HIV2s are about 40% different in nucleotide sequence over the entire genome and differ in presence of some accessory genes. HIV1 and SIVcpz contain two open reading frames, termed *vpr* and *vpu*, whereas HIV2/SIVmac/SIVsm lack the *vpu* gene but contain another gene termed *vpx*. Rather than requiring at least two recent human infections with quite different viruses, the "old virus" hypothesis requires only one cross-species infection of humans.

Current Study

Methods.—We based our phylogenetic analysis of HIVs and SIVs on the relatively slowly evolving *pol* gene and *gag* p24 region of 28 viral isolates obtained from the Los Alamos HIV sequence database (Myers et al., 1993a) and GenBank (NCBI Entrez release 7.0). These sequences are from isolates of both primary HIV types (HIV1 and HIV2), wild-caught African primates (SIVcpz from chimpanzee, *Pan troglodytes*; SIVagms from African green monkeys, *Cercopithecus aethiops* and *C. pygerythrus*; SIVsyk from Sykes' monkey, *Cercopithecus mitis albogularis*; SIVsms from sooty mangabeys, *Cercocebus atys*; and SIVmndgb from mandrill, *Mandrillus sphinx*), captive primates (SIVmacs from *Macaca mulatta*, *M. nemistrina*, and *M. arctoides*), and a domestic cat (FIV from *Felis catus*). We combined regions from both the *pol* and *gag* genes, emphasizing congruence among characters as support for sister-group relationships (Kluge, 1989). Inferred amino acid sequences were aligned using Clustal V (Higgins et al., 1992) with fixed and floating gap penalties set to 10. Those sequence regions that did have gaps were excluded from phylogenetic analyses, leaving unambiguously aligned regions spanning 2,763 nucleotide positions in *pol* and 698 positions in *gag* p24. The amino acid alignment served as a template for aligning the corresponding nucleic acids.

Phylogenetic analysis was conducted on the aligned, edited sequences using PAUP 3.1.1 (Swofford, 1993). Given the large number of terminal taxa and the large size of the data matrices, we used a heuristic search algorithm with the tree bisection-reconnection branch-swapping procedure. Because the heuristic search does not explore all possible topologies to find the shortest tree,

we repeated the search 100 times for each analysis. Each search was initiated using a different randomly constructed starting topology, reducing the possibility that the algorithm would find a local parsimony optimum rather than the universal optimum for a particular data set.

Inaccuracies in phylogenetic analyses stem from an inability to discriminate homologous similarity (due to descent) from homoplastic (convergent or parallel) similarity. Two steps for molecular systematists in making this discrimination are choosing genes that are not saturated with change (having multiple substitutions at individual base positions) and using a data-set-dependent a priori weighting scheme to place greater weight on those characters whose rates of change are relatively slow, because similarities among such characters will tend to include less homoplastic similarity (Mindell and Honeycutt, 1990; Hillis et al., 1993). Toward this end, we calculated the number of third-position transition and transversion changes for codons of *pol* and *gag* p24 DNAs between representative HIV and SIV lineage pairs (Table 1). We expected any homoplastic similarity to be found particularly in third-position transitions. Third positions in codons tend to have faster rates of change because of the greater number of synonymous substitutions that are possible there, relative to first and second positions, and a tendency for transitions to accumulate more rapidly than transversions has long been known (Brown et al., 1982; Graur, 1985).

If DNA sequence characters are saturated with change, the number of inferred changes will not increase as divergence time increases between taxon pairs, i.e., the correspondence between time and increasing amounts of sequence divergence will break down. Among our study taxa, divergences within subsets of HIV1s (excluding HIV1ant70 and HIV1mvp5180) and within HIV2s (excluding HIV2d205 and HIV2uc1) are more recent than divergences among the primary HIV/SIV lineages (HIV1, HIV2, SIVagm, SIVmnd, SIVsyk), which in turn are more recent than the divergence of their common ancestor from FIV (Doolittle et al., 1989; Yokoyama, 1991). Table 1 indicates that *pol* and *gag* p24 third-position transitions are relatively saturated with change; pairwise comparisons with FIV show no more changes than do comparisons among the primary HIV/SIV lineages. Conversely, third-position transversions are relatively unsaturated with change; comparisons with FIV consistently show more changes than do other comparisons. Third-position transition differences from FIV are actually fewer than those of more recent divergences among HIV/SIV lineages, as would be expected when more slowly accumulating transversions begin to overwrite transitions. Comparisons similar to those in Table 1 for codon positions 1 and 2 (data not shown) indicate nonsaturation for both transitions and transversions at those positions. Thus, we gave third-position transitions an a priori weight of zero in our phylogenetic analyses to reduce the confounding effects of nonhomologous similarity.

TABLE 1. Pairwise differences between viral taxa based on 1,153 third codon positions for DNA sequences from the *pol* and *gag* p24 regions combined (see Appendix for sequence sources). Numbers of inferred transitions are below the diagonal, and numbers of inferred transversions are above the diagonal.

	1	2	3	4	5	6	7	8	9
1. HIV1eli	—	9	305	310	307	302	336	320	411
2. HIV1ndk	51	—	308	313	310	303	341	321	412
3. HIV2ben	301	309	—	15	312	297	328	327	414
4. HIV2d194	305	309	146	—	319	300	333	338	417
5. SIVagmtyo	299	312	278	293	—	137	337	299	431
6. SIVagm3	294	295	293	303	328	—	323	282	426
7. SIVsyk	287	284	299	307	287	270	—	338	403
8. SIVmndgb	280	276	290	295	291	288	285	—	404
9. FIV14	251	244	282	282	266	248	290	243	—

The relative support for each node within the minimum-length topology was evaluated using the support index (Bremer, 1988; Källersjö et al., 1992), which denotes the difference in length between the most-parsimonious tree and the shortest tree in which the particular node (clade) is not present. To estimate the support index for a particular clade, we constructed a constraint tree in which the clade is the only re-

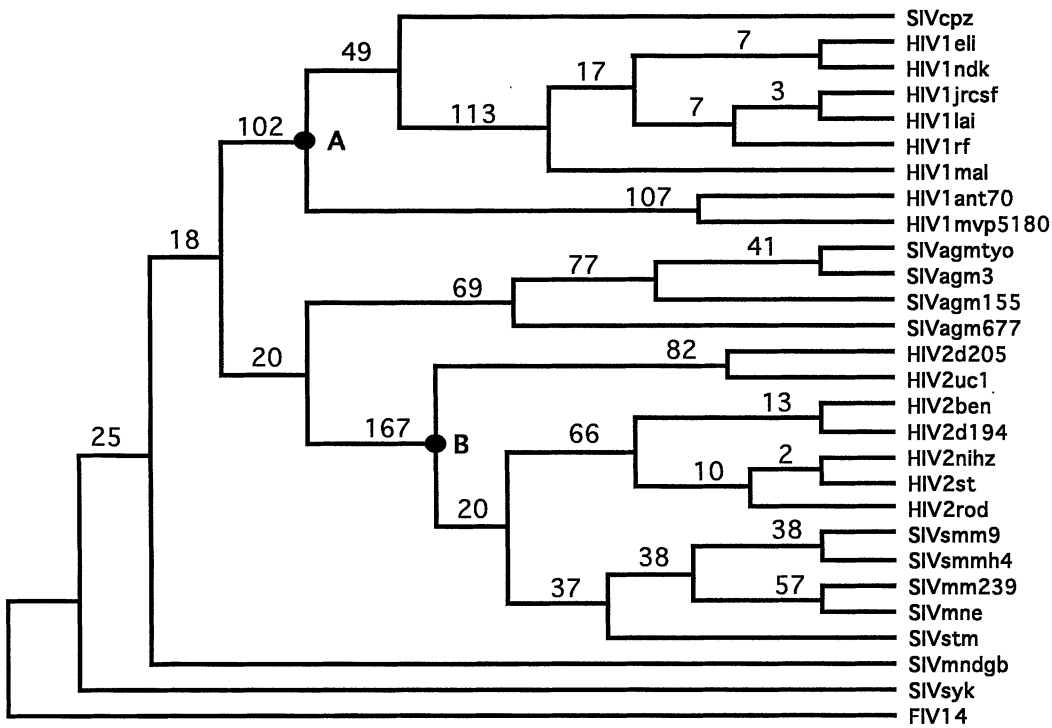

FIGURE 2. Most-parsimonious phylogenetic tree for 27 primate and 1 feline immunodeficiency virus based on the combined DNA sequences from the *pol* (2,763 nucleotide positions) and *gag* p24 (698 nucleotide positions) genes. Third-position transitions were given an a priori weight of zero to reduce homoplastic similarity within the data set (see Table 1). Numbers along nodes are support indices denoting the number of additional steps needed to break the node (Bremer, 1988; Källersjö et al., 1992). The tree length is 6,226 steps, excluding uninformative characters. Nucleotide sequences were obtained from the Los Alamos HIV database and GenBank (see Appendix for accession numbers, abbreviation definitions, and virus host species). Isolates from wild-caught primates include those from chimpanzee (SIVcpz), sooty mangabeys (SIVsmm9, SIVsmmh4), African green monkeys (SIVagm3, SIVagm155, SIVagmtyo, SIVagm677), Sykes' monkey (SIVsyk), and mandrill (SIVmndgb). Isolates from captive primates include those from rhesus macaque (SIVmm239), pig-tailed macaque (SIVmne), and stump-tailed macaque (SIVstm). The feline immunodeficiency virus isolate (FIV14) was treated as an outgroup to the primate immunodeficiency viruses (PIVs). Nodes A and B denote taxa named PIV1 and PIV2, respectively.

solved relationship among the study taxa and then used 10 replicate heuristic searches, with random stepwise addition of taxa, to find the shortest fully resolved topology in which that relationship was not present.

Results.—We found a single most-parsimonious tree in analysis of the combined *pol* and *gag* p24 DNAs, giving third-position transitions zero weight, and we consider this our current best estimate of phylogenetic relationships for the viruses (Fig. 2). Nodes within the tree differ in their degree of support, based on the indices reported. SIVcpz is sister to six HIV1s, and two other HIV1s (HIV1ant70, HIV1mvp5180) are basal to this group. This placement of SIVcpz inside a larger HIV1 clade suggests that there was either one viral host shift from humans to chimp or two host shifts from chimp to humans. Myers et al. (1992:373) suggested that HIV1ant70 "may ultimately be interpreted as yet another SIV form." This interpretation might seem to reduce the likelihood that a human-to-chimp transfer occurred, in that no HIV1 would diverge "prior to" the divergence of SIVcpz. However, interpretation of HIV1ant70 as an SIV (and an aberrant colonist in humans) connotes that HIV1ant70 represents a rare divergent viral lineage in humans. Recent evi-

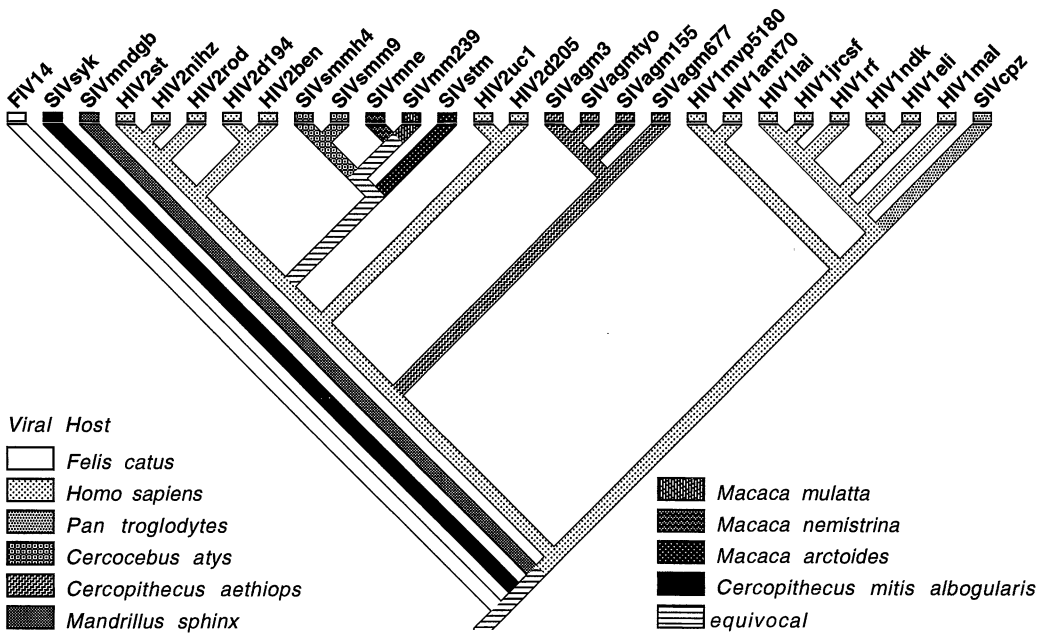

FIGURE 3. Most parsimonious evolution of the character "viral host" based on our most-parsimonious tree topology (presented in Fig. 2). Changes in viral host shown (patterns and shadings of branches) invoke the fewest possible number of shifts. Branches shown as equivocal denote that two or more character states (viral hosts) are possible without altering the minimum number of changes invoked. Because species distinctions among African green monkeys are unclear, we have conservatively listed the four SIVagms shown (here and in the Appendix) as representing one host species, although we note that some might recognize three (SIVagm155, SIVagmtyo, SIVagm3) as being from *Cercopithecus pygerythrus* and one (SIVagm677) as being from *C. aethiops*. Inclusion of two distinct African green monkey species here, however, does not result in any additional changes in the character "viral host" elsewhere in the tree.

dence shows this not to be the case. Nkengasong et al. (1993) found that blood samples from 16 humans from Cameroon and Gabon, Africa, reacted positively with HIV1ant70 peptides in enzyme-linked immunosorbent assays, indicating HIV1ant70 to be endemic among HIV1 seropositive individuals in these two countries and thus more common than previously thought. Further indicating endemism, HIV1mvp5180 (also from a Cameroonian) is the sister to HIV1ant70 at a strongly supported node in our parsimony analysis. Current evidence indicates that human-to-chimp viral transfer is more parsimonious than the reverse (one host species shift versus two; Fig. 3). Publication of sequences from additional chimp SIVs will help resolve this issue.

Our analysis shows HIV2s to be polyphyletic, by placement of two HIV2s (HIV2d205, HIV2uc1) as sister to a clade including HIV2s and SIVs from sooty mangabeys and macaques (Fig. 2), with the latter clade (excluding HIV2d205 and HIV2uc1) being moderately well supported. Like the situation with HIV1s and SIVcpz, the more parsimonious scenario indicates virus transmission from humans (HIV2s) to sooty mangabeys and macaques (Fig. 3). Our tree places SIVmndgb and SIVsyk as basal to the other HIVs and SIVs, although those nodes have relatively moderate levels of support. SIVagms are sister to HIV2s/ SIVsms/ SIVmacs in agreement with most previous analyses but differing from that of Doolittle (1989). Based on current evidence, our topology is preferable because of its incorporation of more character evidence and a data-set-dependent a priori weighting scheme, which reduces effects of homoplastic similarity.

Our tree (Fig. 2) differs from that of Myers et al. (1993b; Fig. 1) in not indicating HIV2 monophyly and in the relative placement of SIVmndgb and SIVsyk. Myers et al. (1993b) used unweighted *gag* p24 region sequences alone, which can be seen to include homoplastic similarity based on our pairwise comparisons (Table 1). Their exclusion of *pol* sequences weakened their analyses, because *pol* includes the most conserved and hence most phylogenetically reliable sequences in the genome. They also used midpoint rooting, which should only be used as a last resort when no suitable outgroup is available. The midpoint method places the basal node for any tree arbitrarily along the longest path connecting any pair of taxa. This placement assumes constant rates of character change across taxa without justification, and differences in rate sufficient to affect placement of the basal node will change sister relationships shown in the tree. Although Myers et al. (1993b) acknowledged the basal position of HIV1ant70 relative to other HIV1s and SIVcpz, they did not include HIV1ant70 (or its sister taxon HIV1mvp5180) in their analysis. This exclusion and their diagnosis of HIV2s as monophyletic, despite numerous analyses contradicting HIV2 monophyly (e.g., Dietrich et al., 1989; Gao et al., 1992; Myers et al., 1992; Barnett et al., 1993) allow Myers et al. (1993b) to consistently favor the "new virus" hypothesis and ignore the alternative "old virus" hypothesis.

Minimum Evolution of the Character "Virus Host"

We used our tree topology (Fig. 2) to infer the most-parsimonious pathway of cross-species infection within the primate immunodeficiency viruses. Host species for each of the 28 viral isolates was coded as the state of a single unordered character, and the minimum number of changes among alternative states were then distributed on the tree (Fig. 3; using MacClade, Maddison and Maddison, 1992). Within the HIV1/SIVcpz clade, human is shown as the ancestral host species. Similarly, within the HIV2/SIVsm/SIVmac/SIVagm clade, human is also shown as the ancestral host species. The basal character state for the entire HIV/SIV clade is equivocal, i.e., two or more different states (virus hosts) could be invoked without altering the number of changes on the overall tree. Recently, Myers and Korbin (1994) included a second SIV from a chimpanzee (SIVcpzant) in their phylogenetic analyses (although this sequence is currently unpublished), and their analysis placed SIVcpzant as sister to the clade including the HIV1s and SIVcpz. Even with inclusion of SIVcpzant in this position on our tree diagnosing change in the character "virus host" (Fig. 3), human remains the most-parsimonious ancestral host for the HIV1/SIVcpz/SIVcpzant clade as well as for the HIV2/SIVsm/SIVmac/SIVagm clade.

Obviously, this analysis does not resolve the sequence of host species shifts, and we make no such claim. Inference from this analysis is confounded by sampling bias, because many more viruses have been sequenced from humans than from any other primate species. Human is shown as the ancestral host for the HIV1/SIVcpz/SIVcpzant and HIV2/SIVsm/SIVmac/SIVagm clades because two of the five most divergent primate immunodeficiency viruses have been isolated from humans, whereas each of the other three divergent viral lineages is unique to a different host species. If, for example, further sampling of SIVs from African green monkeys or mandrills were to uncover taxa within each of those lineages as divergent as HIV1 and HIV2 types are from each other, the character-state changes in Figure 3 would be altered. However, some such sampling has been done for African green monkeys from disparate locales in Africa, and divergences as great as those seen between HIV lineages have not been observed. SIVagms from western Africa (e.g., Senegal) and from eastern Africa (Kenya, Ethiopia; as included in our study), form a monophyletic group (Allan et al., 1991), in contrast to HIV1s from eastern and central Africa and HIV2s from western Africa, which do not form a monophyletic group (Fig. 2). Changing the tree topology in Figure 3, such that SIVagms are basal to the entire HIV1/HIV2 clade, does not change the diagnosis of an-

cestral host species within either the HIV1 or the HIV2 clade. In presenting this analysis, we simply wish to show that the current evidence does not support the "new virus" (new in humans) hypothesis.

Virulent Viruses Are New and Nonvirulent Viruses Are Old: An Oversimplification

It has long been thought that mutualistic associations between parasites and hosts are more stable evolutionarily than are destructive ones (Smith, 1939; Burnet and White, 1972). Parasites that quickly kill their hosts will provide little opportunity for their progeny to successfully colonize new host individuals and, hence, may go extinct. This observation is reflected in a widely claimed (particularly in medical texts) tendency for viruses to evolve toward avirulence and in the general notion that given enough time a state of peaceful coexistence eventually becomes established between any host and parasite (Dubos, 1965).

This assumption has led to an oversimplified prescription that virulent viruses are new and nonvirulent viruses are old. It is becoming increasingly evident, however, that there can be great variation in the timing and direction of virulence change. Just as a virus can change in its effects from pathogenic to benign, it can also change from benign to pathogenic, depending on natural selection and the effects of changing replication rates on the fitness of the virus (Ewald, 1994). As described by May (1993:66),

> There is no generalization [regarding change in virulence for many or most viruses]. The virus may become less virulent, more virulent, or exhibit unchanging virulence; the virus may become less transmissible, more transmissible, or show unchanging transmissibility. All of this depends on the tradeoffs among virulence, transmissibility, and the cost of resistance, which are also constrained by the nature of the host–pathogen association.

Examples of viruses that have shown an increase in virulence over time include influenza A (see Langmuir and Schoenbaum, 1976; Webster, 1993) and myxoma virus (Dwyer et al., 1990; Fenner and Kerr, 1994). Levin and Pimentel (1981) simulated the evolution of a simple system with one host species susceptible to two viral lineages, one of which is more virulent than the other. They found no general trend toward avirulence and that increased virulence may be favored when it increases transmission rate.

In keeping with the older view of virulence, apparent mildness of SIV in sooty mangabeys and African green monkeys has been attributed to an old virus–host association. However, one need not invoke an old association to explain avirulence. The mildness could be attributed to relatively low rates of sexual partner change. Sooty mangabey females apparently have low rates of sexual partner change, restricting copulation to a few males during estrus and not copulating during a prolonged period of maternal care (T. Butynski, pers. comm.). African green monkey females are sexually receptive only seasonally and in groups controlled by a single male (Fedigan and Fedigan, 1988). Thus, potential for rapid spread of SIV through these species appears limited, and viral strains with a rapid replication rate (compromising their host's immune system and health) will have little selective advantage.

Results of laboratory infections of chimps with HIV1 are also inconsistent with the supposition that low virulence denotes a long virus–host association. No AIDS-like disease has been observed among over 100 chimps that have been experimentally infected with HIV1 nor among the minority that have remained infected for 5–10 years (Fultz, 1993; Johnson et al., 1993). Further, in chimps in which HIV1s have become established and have increased in numbers, the capability for successful infection of chimp blood cells has increased (Gendelman et al., 1991; Watanabe et al., 1991), indicating a potential for virulence to increase over time.

The avirulence of SIVs in sooty mangabeys, chimps, mandrills, and other species also remains open to question. A severely ill individual would not last long in nature, as compared with infected but asymptomatic or recovered individuals that could complete normal life spans. For this reason, snapshot seropositivity surveys of existing

populations may underestimate the frequency of infections associated with severe illnesses. A highly virulent, molecularly cloned SIV strain originally from a sooty mangabey (SIVsmpbj; Dewhurst et al., 1990) caused death in experimentally infected sooty mangabeys and macaques (Fultz, 1993), whereas the original parental virus caused a chronic AIDS-like syndrome in macaques and only asymptomatic infection in sooty mangabeys. This finding belies the notions that SIVs in their "natural" host species are exclusively avirulent and that they cannot become more virulent over time. The laboratory transmission that has favored increased virulence in this SIV variant is similar to that proposed for HIV. Rapidly reproducing and severe variants can be maintained if the rapid reproduction provides them with a fitness advantage over more slowly reproducing strains.

The "old virus" hypothesis holds that primitive HIVs may have had low virulence and were maintained in a population that displayed low levels of sexual partner change, perhaps in a rural area. This hypothesis leads to a prediction that some early divergent low-virulence viral strains could still be extant in such populations, and viral isolates with some of these characteristics have been discovered. HIV2d205 and HIV2uc1 represent an early divergent lineage within the HIV2/SIVsm clade and were obtained from asymptomatic individuals from rural Ghana and Ivory Coast, respectively. HIV2uc1 is entirely noncytopathic and readily neutralized by sera from HIV2-infected individuals (Barnett et al., 1993).

Issues in Systematics of Viruses

Conflicting Topologies for Viral Gene Trees May or May Not Indicate Recombination among Viral Lineages

When two or more individual viruses penetrate a particular host cell and begin nucleic acid replication, the potential exists for recombination among the viral genomes due to a replicase enzyme slipping from one viral genome template to another (Coffin, 1979; Hu and Temin, 1990). Recombination among HIV variants occurs in vitro (Clavel et al., 1989) and has been suggested to occur in vivo based on (1) observed viral sequences having a mixture of components from formerly distinct lineages (Howell et al., 1991) and (2) conflicting tree topologies based on phylogenetic analyses of different genes (Li et al., 1988; McClure et al., 1988; Gao et al., 1992; Myers et al., 1993b). Conflicting phylogenies based on different genes, however, may also stem from differential success in phylogenetic analyses. That is, one tree might be accurate whereas the other is not, despite absence of any recombination. In analyzing different genes and different types of substitutions changing at different rates, systematists often find different data sets supporting different trees. This conflict may stem from differential success in distinguishing homologous from homoplastic similarity (distinguishing signal from noise) in the different data sets (Farris, 1983; Swofford and Olsen, 1990; Hillis, 1991; Mindell, 1991). In considering recently diverged taxa, this conflict might also stem from the confounding effects of within-population variation on analyses among higher level taxa (Neigel and Avise, 1986; Avise, 1989). Within-population variation for retroviruses can be extreme, depending on which gene regions are considered (Zarling and Temin, 1976; Holmes et al., 1992). Systematists working on viruses will need to consider these possibilities prior to invoking recombination to explain such conflicts in gene tree topologies.

High Extinction Rates and Sampling Problems

Because of their short generation times, large numbers of progeny, and high mutation rates, viruses have a great capacity for rapid diversification. Consequently, there is also a great capacity for lineage extinction events, which has implications for studies of phylogeny and the history of host shifts. Inclusion of fossil taxa can alter inferred phylogenetic relationships in studies of plants and animals (Doyle and Donoghue, 1987; Gauthier et al., 1988), and the same can be expected in analyses of viruses. More se-

quences from extant viral taxa and, where possible, extinct viral taxa from preserved tissues will help our understanding of the effects of this taxon inclusion/exclusion problem. Peter Houde and colleagues (unpubl. manuscript) at New Mexico State University are working on amplifying and sequencing SIVs from primate museum study skins and, in the process, identifying new host species and minimum dates for host species infection.

Clear distinction must be made between phylogeny of the viruses and the history of their distribution; the two need not be congruent. Divergent lineages such as SIVsyk and SIVmndgb may appear basally in a phylogenetic analysis (as in Fig. 2) without Sykes' monkey or the mandrill being old (early) host species. That basal appearance could be the result of the extinction of lineages from the true early host, which "gave" the virus to current hosts relatively recently, or the result of a lack of sampling from the true early host species. Just as the true phylogeny for any set of taxa is unknowable (unless directly observed) and can only be inferred, the true history of viral host shifts can also only be inferred. For this reason, attempts to determine the natural or ancestral host of a virus will always be susceptible to biases from "unobserved" host shifts, related to high extinction rates for viral lineages and the inevitably small samples available for analysis.

Rate Heterogeneity

RNA viruses such as the primate immunodeficiency viruses, with base substitution rates averaging 10^{-3} per site per year, often have rates of evolution exceeding that of their eukaryotic host species by 1 million-fold or more (Holland, 1992). This is a result of the high error rate of the virus-encoded reverse transcriptase and the lack of misincorporation repair mechanisms. Although this rapid rate of viral sequence change is not qualitatively different from that encountered by systematists working on other taxa, there are several sources of rate variability among viruses that are not currently recognized in other taxa. Retroviruses undergo replication involving three different enzymes with variable error rates. In the viral stage (in the host cell cytoplasm), retroviral RNA is transcribed into retroviral DNA by reverse transcriptase, which has a high error rate. In the proviral stage (in the host cell nucleus), retroviral DNA is replicated by the host cell's DNA polymerase, which is less error prone and entails efficient mutation repair mechanisms. Subsequently, the proviral DNA is transcribed back into RNA by the host cell's RNA polymerase. The error rate for cellular RNA polymerase is not well known, although it may be similar to that of reverse transcriptase (Coffin, 1991). Thus, there is the potential for closely related viral lineages to differ in their rates of change due to different amounts of high-error (reverse transcriptase and cellular RNA polymerase) and relatively low-error (cellular DNA polymerase) replication. These differences will tend to vary with changing virulence; low virulence entails longer proviral times and fewer replication cycles, and high virulence entails greater amounts of low-fidelity reverse transcription. A further consequence of the proviral stage is the opportunity for recombination with cellular genes and the possible addition of new sequences into the retrovirus genome (Bishop and Varmus, 1985).

We expect that rates of retroviral change may vary depending upon the particular host species infected, given that different animal species may show different rates of molecular sequence evolution (e.g., Britten, 1986; Li and Tanimura, 1987; Avise et al., 1992; Martin and Palumbi, 1993) and that retroviruses use the host's replication machinery. Rates might also vary depending on the particular cell type infected, as suggested by correlation between rates of sequence change and metabolic rate (rate of oxygen metabolism) and by differences in metabolic rate for different cell and tissue types. Underlying the correlation with metabolic rate is apparent DNA damage due to oxygen-derived free radicals (Joenje, 1989; Shigenaga et al., 1989). Oxidative damage potentially influences rates of sequence evolution across all taxa; however,

the generally fast rate of retroviral evolution accentuates these and other effects to a greater degree than is seen in other organisms.

Viral sequences also show patterns of rate heterogeneity correlated with codon position and transition/transversion differences, as seen in other organisms (Graur, 1985; Table 1). We have sought to account for the effects of some of these in our current analyses with an a priori weighting scheme. The effect on phylogeny of other rate heterogeneity sources mentioned (three different replication enzymes, host- and cell-specific effects) are poorly known at present, although they are potentially significant. In light of the fast pace of primary sequence evolution and subsequent low levels of sequence similarity among many viral taxa, the more slowly evolving features of secondary and tertiary structures for encoded proteins may prove useful for alignment and phylogenetic analyses in the future (see Johnson et al., 1990; Eickbush, 1994).

Estimates of lineage divergence times assume rate constancy over time and will be distorted to the extent that rate heterogeneity exists for the characters analyzed. Not surprisingly, this distortion has given rise to incongruent estimates by different researchers. Estimates for divergence time between HIV1s and HIV2s range from 40 (Smith et al., 1988) to 600–1,200 years ago (Eigen and Nieselt-Struwe, 1990).

Naming Virus Clades Rather Than Grades

The names used for primate immunodeficiency virus taxa have been based on the host species in which the viruses have been found. Thus, these names represent viral grades based on virus distribution. Named grades are less desirable than named clades, given that the primary purpose of taxonomy is to communicate results of evolutionary history (phylogenetic analysis) using a system of names. As more viral taxa become known and are added to phylogenetic analyses, viral taxonomy can be revised to provide a more accurate history of their evolution. Such a revision can discourage misconceptions or premature conclusions regarding lineage origins. For example, the association by name of HIV1s and HIV2s suggests (to systematists) a common origin for them to the exclusion of other immunodeficiency viruses, but (as discussed above) this appears not to be the case. Similarly, the taxon SIV gives the unsupported impression that all SIVs are more closely related to each other than they are to various HIVs. De Queiroz and Gauthier (1992) described useful conventions for naming taxa, the most basic of which is that all names refer to clades.

We can recognize the clades in Figure 2 as taxa. The clade that is descendent from the hypothetical common ancestor at node A in Figure 2 includes all the known HIV1s and SIVcpz and can be called primate immunodeficiency virus 1 (PIV1). The clade that is descendent from the hypothetical common ancestor at node B in Figure 2 includes all the known HIV2s and SIVs from sooty mangabeys and macaques and can be called PIV2. Other taxa may be recognized in a similar fashion as the need arises. Members of the taxon PIV1 have an apparent synapomorphy in the presence of the *vpu* accessory gene, whereas members of PIV2 uniquely possess the accessory genes *vpr* and *vpx* in combination (Gibbs and Desrosiers, 1994). In a nonphylogenetic taxonomy, such characters might have been used to define taxa. However, in our proposed phylogenetic taxonomy, such characters are used in diagnosing clades but not in defining them (determining inclusion or exclusion of species or taxa). Rather, taxon names are defined in terms of common ancestry and relationship.

Conclusions

Evidence currently available does not support the popular view (the "new virus" hypothesis) that HIVs (our PIVs, Fig. 2) have recently colonized humans and that PIVs in humans are recent descendants from one or another of the PIV lineages known from nonhumans. Phylogenetic trees show only sister relationships for extant taxa, not ancestor–descendant relationships for extant taxa. Our phylogenetic hypothesis and a parsimony criterion to estimate the fewest

number of host species shifts (i.e., to diagnose changes in the character "viral host") indicate that humans are the ancestral host species for a clade including SIVcpz (from chimpanzee) and for a clade including SIVsms (from sooty mangabeys). However, in light of potential sampling biases, we specifically do not claim that this analysis resolves the issue of ancestral host. Our point is to show that current evidence does not support the "new virus" hypothesis. Support for the "new virus" hypothesis then devolves to unjustified assumptions that (1) pre-1959 human blood samples testing negative for PIV presence successfully represent all human populations and demes potentially harboring PIVs and (2) new viruses are virulent and old viruses are mild. Small human populations with dormant PIVs may readily have been missed by limited sampling, and the assumption that new viruses are virulent and old viruses are mild ignores the ability of natural selection to affect an increase, a decrease, or no change in virulence over time. Even if the decreasing virulence assumption were valid, inferred newness of PIV infection of humans is contradicted by discovery of noncytopathic HIV2uc1 and relatively low virulence (longer latency and asymptomatic periods) of PIV2s in rural human populations having relatively low rates of sexual contact among individuals.

Retroviral evolution challenges systematists with a variety of distinctive and potentially confounding features, including (1) extremely fast rates of molecular sequence evolution (due to short generation times, large numbers of progeny, and low fidelity replication), (2) evolutionary rate heterogeneity within and among virus sequences (due to potential host-specific and cell-type-specific rate differences and variable use of three different replication enzymes having variable error rates), and (3) potential for genetic recombination among different lineages infecting the same cell, complicating character homology determinations. Improved understanding of these features and greater sampling of primate host species will enhance future studies of immunodeficiency virus phylogeny and may entail revision of current hypotheses of relationship.

Acknowledgments

We thank J. J. Bull, M. J. Donoghue, D. M. Hillis, T. D. Kocher, M. M. Miyamoto, G. Myers, and T. W. Scott for valuable discussion and comments on various drafts of this manuscript. This work was supported by the National Science Foundation (BSR-9019669 to D.P.M.) and an Alfred P. Sloan Postdoctoral Fellowship (to J.W.S.).

References

Allan, J. S., M. Short, M. E. Taylor, S. Su, V. M. Hirsch, P. R. Johnson, G. M. Shaw, and B. H. Hahn. 1991. Species–specific diversity among simian immunodeficiency viruses from African green monkeys. J. Virol. 65:2816–2828.

Avise, J. C. 1989. Gene trees and organismal histories: A phylogenetic approach to population biology. Evolution 43:1192–1208.

Avise, J. C., B. W. Bowen, T. A. Lamb, B. Meylan, and E. Bermingham. 1992. Mitochondrial DNA evolution at a turtle's pace: Evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Mol. Biol. Evol. 9:457–473.

Barnett, S. W., M. Quiroga, A. Werner, D. Dina, and J. A. Levy. 1993. Distinguishing features of an infectious molecular clone of the highly divergent and noncytopathic human immunodeficiency virus type 2 UC1 strain. J. Virol. 67:1006–1014.

Bishop, J. M., and H. Varmus. 1985. Function and origins of retroviral transforming genes. Pages 249–356 *in* RNA tumor viruses: Molecular biology of tumor viruses, 2nd edition, Volume 2 [supplements and appendices] (R. Weiss, N. Teich, H. Varmus, and J. Coffin, eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Bremer, K. 1988. The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

Britten, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1393–1398.

Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J. Mol. Evol. 18:225–239.

Burnet, F. M., and D. O. White. 1972. Natural history of infectious disease, 4th edition. Cambridge Univ. Press, Cambridge, England.

Clavel, F., M. D. Hoggan, J. R. Willey, K. Strebel, M. A. Martin, and R. Repaske. 1989. Genetic recombination of human immunodeficiency virus. J. Virol. 63:1455–1459.

Coffin, J. M. 1979. Structure, replication, and recombination of retrovirus genomes; some unifying hypotheses. J. Gen. Virol. 42:1–26.

Coffin, J. M. 1991. Retroviridae and their replication. Pages 645–708 *in* Fundamental virology (B. N. Fields and D. M. Knipe, eds.). Raven Press, New York.

DE QUEIROZ, K., AND J. GAUTHIER. 1992. Phylogenetic taxonomy. Annu. Rev. Ecol. Syst. 23:449–480.

DEWHURST, S., J. E. EMBRETSON, D. C. ANDERSON, J. I. MULLINS, AND P. N. FULTZ. 1990. Sequence analysis and acute pathogenicity of molecularly cloned SIV smm-PBj14. Nature 345:636–640.

DIETRICH, U., M. ADAMSKI, R. KREUTZ, A. SEIPP, H. KUHNEL, AND H. RUBSAMEN-WAIGMANN. 1989. A highly divergent HIV-2 related isolate. Nature 342:948–950.

DOOLITTLE, R. F. 1989. Immunodeficiency viruses: The simian–human connection. Nature 339:338–339.

DOOLITTLE, R. F., D.-F. FENG. M. S. JOHNSON, AND M. A. MCCLURE. 1989. Origins and evolutionary relationships of retroviruses. Q. Rev. Biol. 64:1–30.

DOYLE, J. A., AND M. J. DONOGHUE. 1987. The importance of fossils in elucidating seed plant phylogeny and macroevolution. Rev. Palaeobot. Palynol. 50:63–95.

DUBOS, R. 1965. Man adapting. Yale Univ. Press, New Haven, Connecticut.

DWYER, G., S. A. LEVIN, AND L. BUTTEL. 1990. A simulation model of the population dynamics and evolution of myxomatosis. Ecol. Monogr. 60:423–447.

EICKBUSH, T. H. 1994. Origin and evolutionary relationships of retroelements. Pages 121–157 *in* The evolutionary biology of viruses (S. S. Morse, ed.). Raven Press, New York.

EIGEN, M., AND K. NIESELT-STRUWE. 1990. How old is the immunodeficiency virus? AIDS 4(suppl. 1):s85–s93.

EWALD, P. W. 1991. Transmission modes and the evolution of virulence, with special reference to cholera, influenza and AIDS. Human Nature 2:1–30.

EWALD, P.W. 1994. The evolution of infectious disease. Oxford Univ. Press, New York.

FARRIS, J. S. 1983. The logical basis of phylogenetic analysis. Pages 7–36 *in* Advances in cladistics (N. W. Platnick and V. A. Funk, eds.). Columbia Univ. Press, New York.

FEDIGAN, L., AND L. M. FEDIGAN. 1988. *Cercopithecus aëthiops*: a review of field studies. Pages 389–411 *in* A primate radiation: Evolutionary biology of the African guenons (A. Gautier-Hion, F. Bourliere, J.-P. Gautier, and J. Kingdon, eds.). Cambridge Univ. Press, Cambridge, England.

FENNER, F., AND P. J. KERR. 1994. Evolution of the poxviruses, including the coevolution of virus and host in myxomatosis. Pages 273–292 *in* The evolutionary biology of viruses (S. S. Morse, ed.). Raven Press, New York.

FOX, C. 1992. Possible origins of AIDS. Science 256:1259–1260.

FULTZ, P. N. 1993. Nonhuman primate models for AIDS. Clin. Infect. Dis. 17(suppl. 1):S230–S235.

GAO, F., L. YUE, A. T. WHITE, P. G. PAPPAS, J. BARCHUE, A. P. HANSON, B. M. GREENE, P. M. SHARP, G. M. SHAW, AND B. H. HAHN. 1992. Human infection by genetically diverse SIVsm-related HIV-2 in West Africa. Nature 358:495–499.

GAUTHIER, J., A. G. KLUGE, AND T. ROWE. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4:105–209.

GENDELMAN, H. E., G. D. EHRLICH, L. M. BACA, S. CONLEY, J. RIBAS, D. C. KALTER, M. S. MELTZER, B. J. POIESZ, AND P. NARA. 1991. The inability of human immunodeficiency virus to infect chimpanzee monocytes can be overcome by serial viral passage in vivo. J. Virol. 65:3853–3863.

GIBBS, J. S., AND R. C. DESROSIERS. 1994. Auxiliary proteins of the primate immunodeficiency viruses. Pages 137–158 *in* Human retroviruses (B. R. Cullen, ed.). Oxford Univ. Press, New York.

GRAUR, D. 1985. Pattern of nucleotide substitution and the extent of purifying selection in retroviruses. J. Mol. Evol. 21:221–231.

GRMEK, M. D. 1990. History of AIDS: Emergence and origin of a modern pandemic. Princeton Univ. Press, Princeton, New Jersey.

HAHN, B. H. 1990. Biologically unique SIV-like HIV-2 variants in healthy West African individuals. Pages 31–38 *in* Retroviruses of human A.I.D.S. and related animal diseases (M. Girard and L. Valette, eds.). Fondation Marcel Mérieux, Lyon, France.

HAHN, B. H. 1994. Viral genes and their products. Pages 21–43 *in* Textbook of AIDS medicine (S. Broder, T. Merigan, and D. Bolognesi, eds.). Williams & Wilkins, Baltimore, Maryland.

HIGGINS, D., A. J. BLEASBY, AND R. FUCHS. 1992. Clustal V: Improved software for multiple sequence alignment. CABIOS 8:189–191.

HILLIS, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. Pages 278–294 *in* Phylogenetic analysis of DNA sequences (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.

HILLIS, D. M., M. W. ALLARD, AND M. M. MIYAMOTO. 1993. Analysis of DNA sequence data: Phylogenetic inference. Methods Enzymol. 224:456–487.

HIRSCH, V. M., G. A. DAPOLITO, S. GOLDSTEIN, H. MCCLURE, P. EMAU, P. N. FULTZ, M. ISAHAKIA, R. LENROOT, G. MYERS, AND P. R. JOHNSON. 1993. A distinct African lentivirus from Sykes' monkeys. J. Virol. 67: 1517–1528.

HIRSCH, V. M., R. A. OLMSTED, M. MURPHEY-CORB, R. H. PURCELL, AND P. R. JOHNSON. 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature 339:389–392.

HOLLAND, J. J. (ed.). 1992. Genetic diversity of RNA viruses. Current topics in microbiology and immunology, No. 176. Springer-Verlag, New York.

HOLMES, E. C., L. Q. ZHANG, P. SIMMONDS, C. A. LUDLAM, AND A. J. LEIGH BROWN. 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. Proc. Natl. Acad. Sci. USA 89:4835–4839.

HOWELL, R. M., J. E. FITZGIBBON, N. NOE, Z. REN, D. J. GOCKE, T. A. SCHWARTZER, AND D. T. DUBIN. 1991. In vivo sequence variation of the human immunodeficiency virus type 1 *env* gene: Evidence for recombination among variants found in a single individual. AIDS Res. Hum. Retroviruses 7:869–876.

HU, W. S., AND H. M. TEMIN. 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: Pseudoploidy and high rate of genetic recombination. Proc. Natl. Acad. Sci. USA 86: 9253–9257.

HUET, T., R. CHEYNIER, A. MEYERHANS, G. ROELANTS, AND S. WAIN-HOBSON. 1990. Genetic organization of a

chimpanzee lentivirus related to HIV-1. Nature 345:356–358.

JOENJE, H. 1989. Genetic toxicology of oxygen. Mutat. Res. 219:193–208.

JOHNSON, B. K., G. A. STONE, M. S. GODEC, D. M. ASHER, D. C. GAJDUSEK, AND C. J. GIBBS, JR. 1993. Long-term observations of human immunodeficiency virus-infected chimpanzees. AIDS Res. Hum. Genet. 9: 375–378.

JOHNSON, M. S., A. SALI, AND T. L. BLUNDELL. 1990. Phylogenetic relationships from three-dimensional protein structures. Methods Enzymol. 183:670–690.

KÄLLERSJÖ, M., J. S. FARRIS, A. G. KLUGE, AND C. BULT. 1992. Skewness and permutation. Cladistics 8: 275–287.

KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38:7–25.

LANGMUIR, A. D., AND S. C. SCHOENBAUM. 1976. The epidemiology of influenza. Hosp. Pract. 11:49–56.

LEVIN, S. A., AND D. PIMENTEL. 1981. Selection for intermediate rates of increase in parasite–host systems. Am. Nat. 117:308–315.

LI, W.-H., T. GOJOBORI, AND P. M. SHARP. 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. 5:313–330.

LI, W.-H., AND M. TANIMURA. 1987. The molecular clock runs more slowly in man than in apes. Nature 326:93–96.

MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade: Analysis of phylogeny and character evolution. Sinauer, Sunderland, Massachusetts.

MARTIN, A. P., AND S. R. PALUMBI. 1993. Body size, metabolic rate, generation time, and molecular clock. Proc. Natl. Acad. Sci. USA 90:4087–4091.

MAY, R. M. 1993. Ecology and evolution of host–virus associations. Pages 58–68 *in* Emerging viruses (S. S. Morse, ed.). Oxford Univ. Press, New York.

MCCLURE, M. 1990. Where did the AIDS virus come from? New Sci. 126(1723):54–57.

MCCLURE, M. A., M. S. JOHNSON, D.-F. FENG, AND R. F. DOOLITTLE. 1988. Sequence comparisons of retroviral proteins: Relative rates of change and general phylogeny. Proc. Natl. Acad. Sci. USA 85:2469–2473.

MINDELL, D. P. 1991. Aligning DNA sequences: Homology and phylogenetic weighting. Pages 73–89 *in* Phylogenetic analysis of DNA sequences (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.

MINDELL, D. P., AND R. L. HONEYCUTT. 1990. Ribosomal RNA in vertebrates: Evolution and phylogenetic applications. Annu. Rev. Ecol. Syst. 21:541–566.

MONTAGNIER, L. 1985. Lymphadenopathy-associated virus: From molecular biology to pathogenicity. Ann. Intern. Med. 103:689–693.

MYERS, G., AND B. KORBER. 1994. The future of human immunodeficiency virus. Pages 211–232 *in* The evolutionary biology of viruses (S. S. Morse, ed.). Raven Press, New York.

MYERS, G., B. KORBER, S. WAIN-HOBSON, R. F. SMITH, AND G. N. PAVLAKIS (eds.). 1993a. Human retroviruses and AIDS 1993. Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico.

MYERS, G., D. MACINNES, AND B. KORBER. 1992. The emergence of simian/human immunodeficiency viruses. AIDS Res. Hum. Retroviruses 8:373–386.

MYERS, G., K. MACINNES, AND L. MYERS. 1993b. Phylogenetic moments in the AIDS epidemic. Pages 120–137 *in* Emerging viruses (S. S. Morse, ed.). Oxford Univ. Press, New York.

NEIGEL, J. E., AND J. C. AVISE. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pages 515–534 *in* Evolutionary processes and theory (E. Nevo and S. Karlin, eds.). Academic Press, New York.

NKENGASONG, J. N., M. PEETERS, M. VAN DEN HAESEVELDE, S. S. MUSI, B. WILLEMS, P. M. NDUMBE, E. DELAPORTE, J. L. PERRET, P. PIOT, AND G. VAN DEN GROEN. 1993. Antigenic evidence of the presence of the aberrant HIV1ant70 virus in Cameroon and Gabon. AIDS 7:1536–1537.

SHIGENAGA, M. K., C. J. GIMENCO, AND B. N. AMES. 1989. Urinary 8-hydroxy-2′-deoxyguanosine as a biological marker in in vivo oxidative DNA damage. Proc. Natl. Acad. Sci. USA 86:9697–9701.

SHOPE, R. E., AND A. S. EVANS. 1993. Assessing geographic and transport factors, and recognition of new viruses. Pages 109–119 *in* Emerging viruses (S. S. Morse, ed.). Oxford Univ. Press, New York.

SMITH, T. 1939. Parasitism and disease. Princeton Univ. Press, Princeton, New Jersey.

SMITH, T. F., A. SRINIVASAN, G. SCHOCHETMAN, M. MARCUS, AND G. MYERS. 1988. The phylogenetic history of immunodeficiency viruses. Nature 333:573–575.

STINE, G. J. 1993. AIDS update 1993. Prentice Hall, Englewood Cliffs, New Jersey.

SWOFFORD, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.

SWOFFORD, D. L., AND G. J. OLSEN. 1990. Phylogeny reconstruction. Pages 411–501 *in* Molecular systematics (D. M. Hillis and C. Moritz, eds.). Sinauer, Sunderland, Massachusetts.

TEMIN, H. M. 1993. The high rate of retrovirus variation results in rapid evolution. Pages 219–225 *in* Emerging viruses (S. S. Morse, ed.). Oxford Univ. Press, New York.

WATANABE, M., D. J. RINGLER, P. N. FULTZ, J. J. MACKEY, J. E. BOYSON, C. G. LEVINE, AND N. L. LETVIN. 1991. A chimpanzee-passaged human immunodeficiency virus isolate is cytopathic for chimpanzee cells but does not induce disease. J. Virol. 65:3344–3348.

WEBSTER, R. G. 1993. Influenza. Pages 37–45 *in* Emerging viruses (S. S. Morse, ed.). Oxford Univ. Press, New York.

YOKOYAMA, S. 1991. Molecular evolution of human immunodeficiency viruses and related retroviruses. Pages 96–111 *in* Evolution at the molecular level (S. K. Selander, A. G. Clark, and T. S. Whittam, eds.). Sinauer, Sunderland, Massachusetts.

ZARLING, D. A., AND H. M. TEMIN. 1976. High spontaneous mutation rate of an avian sarcoma virus. J. Virol. 17:74–84.

Received 4 February 1994; accepted 17 October 1994

Note added in proof.—Findings published since final acceptance of our paper contribute to the picture of HIV and SIV phylogeny. Peeters et al. (Peeters, M., W. Janssens, K. Fransen, J. Brandful, L. Heyndrickx, K. Koffi, E. Delaporte, P. Piot, G.-M. Gershy-Damet, and G. van der Groen. 1994. Isolation of simian immonudeficiency viruses from two sooty mangabeys in Côte d'Ivoire: Virological and genetic characterization and relationship to other HIV type 2 and SIVsm/mac strains. AIDS Res. Hum . Retroviruses 10:1289–1294) described a second group of SIVs from sooty mangabeys (SIVsmc12, SIVsmc18) that appear to branch off between the earliest diverging HIV2 clade and the SIV sooty mangabey/macaque clade. This suggests even more interspecific transfers than shown in our Figure 3, although inclusion of the new SIVsms in the Figure 3 analysis does not change *Homo* as the postulated early host subsequent to the branching of SIVmndgb. Sharp et al. (Sharp, P. M., D. L. Robertson, F. Gao, and B. H. Hahn. 1994. Origins and diversity of human inmmunodeficiency viruses. AIDS 8:s27–s42) independently proposed a classificatory change similar to our PIV1 and PIV2, emphasizing phylogenetic relatedness, although they suggested the term PLV (primate lentiviruses).

APPENDIX. Immunodeficiency virus abbreviations, host species, and database sequence accession numbers.

Virus abbreviation	Host species	Accession no.
FIV14	*Felis catus*	M25381, M25729
HIV1ant70	*Homo sapiens*	L20587, M31171
HIV1eli	*H. sapiens*	K03454
HIV1ndk	*H. sapiens*	M27323
HIV1jrcsf	*H. sapiens*	M38429
HIV1lai	*H. sapiens*	K02013
HIV1mal	*H. sapiens*	K03456
HIV1mn	*H. sapiens*	M17449
HIV1mvp5180	*H. sapiens*	L20571
HIV1rf	*H. sapiens*	M17451, M12508
HIV2ben	*H. sapiens*	M30502
HIV2d194	*H. sapiens*	J04542, X52223
HIV2d205	*H. sapiens*	X16109, X61240
HIV2nihz	*H. sapiens*	J03654
HIV2rod	*H. sapiens*	M15390
HIV2st	*H. sapiens*	M31113
HIV2uc1	*H. sapiens*	L07625
SIVagm3	*Cercopithecus aethiops*	M30931
SIVagm9	*Cercopithecus tantalus*	L19254
SIVagm40	*Cercopithecus tantalus*	L19252
SIVagm49	*Cercopithecus tantalus*	L19253
SIVagm155	*Cercopithecus aethiops*	M29975
SIVagm677	*Cercopithecus aethiops*	M66437
SIVagm692	*Cercopithecus pygerythrus*	M29974
SIVagmtyo	*Cercopithecus aethiops*	X07805
SIVcpz	*Pan troglodytes*	X52154
SIVmm142	*Macaca mulatta*	M16403, Y00277
SIVmm239	*Macaca mulatta*	M33262
SIVmm251	*Macaca mulatta*	M19499, M15897
SIVmndgb	*Mandrillus sphinx*	M27470, X15781
SIVmne	*Macaca nemistrina*	M32741
SIVsmmh4	*Cercocebus atys*	X14307
SIVsmm9	*Cercocebus atys*	M80194
SIVsmpbj	*Cercocebus atys*	M31325
SIVstm	*Macaca arctoides*	M83293
SIVsyk	*Cercocebus mitis albogularis*	L06042