FOIA Number: 2018-0756-F

FOIA MARKER

This is not a textual record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

Collection/Record Group: Clinton Presidential Records

Subgroup/Office of Origin: National AIDS Policy Office

Series/Staff Member: Kristine Gebbie

Subseries:

OA/ID Number: 3773

FolderID:

Folder Title:
Albuquerque, New Mexico Trip 3/25-27/94 [2]

Stack:	Row:	Section:	Shelf:	Position:
S	66	5	4	2

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

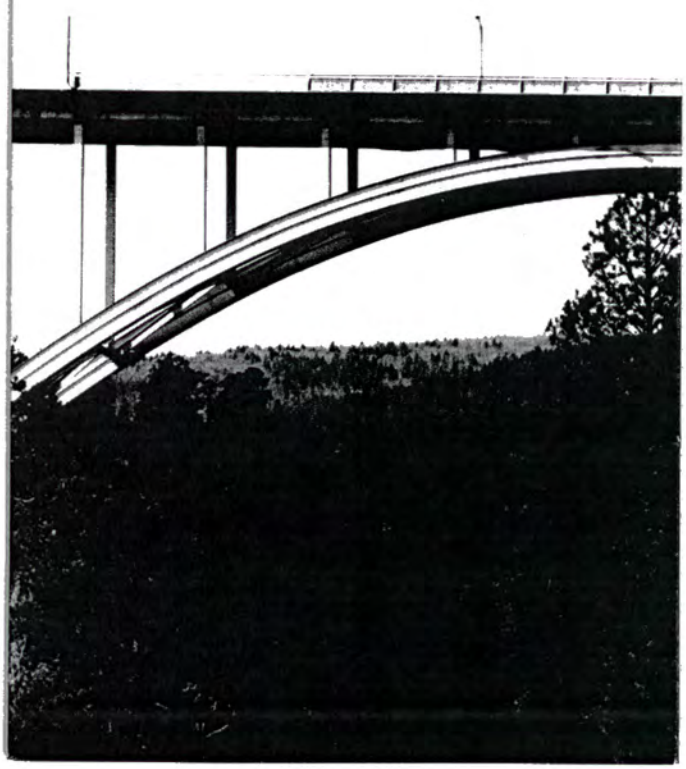
Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.

Visitor's Guide
Los Alamos National Laboratory

6pgs

VISITOR'S GUIDE

Los Alamos National Laboratory



The Evolutionary Biology of Viruses,
edited by Stephen S. Morse.
Raven Press, Ltd., New York © 1994.

10

The Future of Human Immunodeficiency Virus

Gerald Myers and Bette Korber

HIV Sequence Database and Analysis Project, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

"All is flux."
—Heraclitus

The phylogenetic processes of most viruses are slow compared with the rate at which their diseases spread. Even when the genetic variation of a pathogenic virus is unusually brisk, making vaccine modifications a perennial challenge, the dimensions of the disease typically remain unchanged (1). Human immunodeficiency virus (HIV), which causes acquired immunodeficiency syndrome (AIDS) in humans, is an exceptional case: as a result of a recent ecological breakthrough it appears to be engaged in a process of "fast-forward" evolution. This chapter will argue, on the basis of genetic indications, that the future of HIV will be characterized not only by continued rapid genetic variation, but also by a significant broadening of the clinical spectrum of AIDS.

Evolutionary biology is on safer ground when it inquires into the past, rather than the future. Although some unanswered questions remain about the "deep" history of the AIDS viruses, our understanding of their recent emergence has grown swiftly since the publication of the first HIV genetic sequences in 1985 (1–11; see also chap. 10). Although we will not systematically review that body of information, the first part of this chapter will consider certain aspects of the speciation and molecular epidemiology of HIV that we deem most predictive. What is known of its past suggests that HIV is in a singularly threatening state of evolutionary disequilibrium.

In the second part of the chapter, some quantitative features of HIV variation will be examined. First, the global rate of HIV variation will be discussed. Then we will turn to an extraordinary characteristic of all known primate immunodeficiency viruses: they are highly variable and also *complex*. (By complex we mean that they are able to regulate their own gene expression; 11). These would seem to be opposing evolutionary tendencies: Some viruses, such as the human T-cell lymphotropic retroviruses (HTLV), are complex, but manifest comparatively lower rates of genetic change; other RNA viruses mutate rapidly, but their near-total dependence on intri-

cate aspects of host cell machinery places considerable constraint on their degrees of evolutionary freedom. The primate immunodeficiency viruses are degenerate in their complexity. They display inordinately high frequencies of nonsynonymous substitutions, not simply, as we might expect, in their envelope protein-coding sequences, but also in their regulatory and accessory genes. This motley evolutionary posture, coupled with the familiar collection of retroviral adaptive strategies–integration, legitimate and illegitimate recombination, transduction, pseudotyping, molecular mimicry and parall[illegible]s crucial pathogenic implications.

In the third part of the [illegible], we will consider a few of the many different hypotheses about AIDS pathogenesis. We will then explore the association of different pathogenic characteristics with different genotypic patterns in both HIVs and simian immunodeficiency viruses (SIVs), focusing attention on the now well-studied HIV V3 loop, a small peptide in the viral coat protein that has an established association with HIV phenotype. On the basis of a *phenetic* analysis of HIV V3 loop variation, we will argue that the seemingly untrammeled diversification of HIV constitutes a significant factor in its pathogenic potential.

THE RECENT EMERGENCE OF HIV

Simian Origin of HIV: An Evolutionary Turning Point

There is strong evidence that HIVs and SIVs have naturally evolved from an unknown precusor within the highly diverse subfamily of lentiretroviruses, found also in ungulates and felids (12). The worldwide prevalence of these retroviruses suggests that they are old (13), but there is as yet no consensus about how old they are, especially in relation to the divergence of HIVs and SIVs. The European and South African visna viruses, which infect domesticated sheep, diverged from one another about 50 years ago, and one group of researchers has, on the basis of this fact, hypothesized a minimum "lookback" time for lentiviral radiation of merely 400 years (14). Eigen's calculations, in contrast, place the oldest common ancestor for known human and simian lentiviruses back at least 600 to 1,200 years (5), and it has been argued that even this seems too short a time for the dissemination of the felid lentiviruses (13). What is more, the African green monkey lentivirus (SIVagm), which may well be the distant predecessor of HIVs, is believed to have been present before the allopatric speciation of host guenons several thousand years ago; the four extant species of green monkeys dispersed throughout Africa harbor four distinguishable subtypes of SIV (15).

These reckonings—hundreds of years, thousands of years—are arguably short spans even on a microbial evolutionary scale. Yet, precisely because so much molecular genetic activity has gone on in the short evolutionary stretch of the lentiviruses, the deeper branches of the lentiviral tree are enveloped in darkness (9). This indeterminacy will not, however, be a consideration of any weight for our purposes, because our focus will be on the primate lentiviruses, particularly the recent emergence of HIVs from SIVs, to which we will now turn.

Three historical moments, each traceable to the 1970s, mark the onset of AIDS in the world. These three events are the epidemic of HIV-1 disease, the separate epidemic of HIV-2 disease and the outbreak (epizootic) of AIDS-like disease in captive macaques, animals not found to be infected in the wild. Two earlier incidents of apparent HIV-1 infection have been documented retrospectively; the one referred to as the "Manchester sailor" case (16), involving a seropositive sample from 1959, is now undergoing molecular investigation. The preponderance of HIV-1 AIDS viewed from the dual standpoint of clinical and seroepidemiological records, however, came forward in Africa and the United States in the 1970s (2,10). What began as a handful of cases known anecdotally in the 1960s and 1970s has now resulted in the estimated infection of more than 10 million individuals: HIV-1 alone is expected to infect more than 40 million people by the year 2000 (17,18).

HIV-2 AIDS, which, until recently, has largely been associated with people of West Africa, does not have an equally well-documented seroepidemiologic record (10). However, the evidence of early cases of HIV-2 dissemination to India (19), Spain (20), Brazil (21), and even the near-epicentric Cameroon (22) is suggestive of a new and separate AIDS pandemic. The HIV-2 epidemic in India appears to be aggressive compared with the current low level of infection elsewhere outside of western Africa (19).

Finally, sooty mangabey viruses (SIVsmm) that are virtually indistinguishable from HIV-2s (see later) contributed to a devastating epizootic in the mid-1970s. The SIVsmm cross-infected stump-tailed macaques housed with the viruses' natural hosts, mangabeys from West Africa, first at the California Regional Primate Center in the 1960s, then subsequently at other United States primate centers (23,24).

This historical background, together with molecular evidence (discussed in a later section), has led AIDS researchers to conclude that three independent events of cross-species viral transmission and subsequent dissemination in the recipient species occurred at about the same time. Explanations for this striking improbability have focused on accidents and social conditions: (a) increased handling of animals associated with the sharp increase of animal exportation from Africa, (b) the widespread use of needles in association with human and captive monkey vaccination programs, and (c) urbanization of Africa (2,3,10,15,17,23). We may never know which of these contributing factors was most responsible for the origin of AIDS.

To illustrate the kind of molecular epidemiological evidence (reviewed in 9–11) that has supported the hypothesis of a recent simian origin of AIDS, one need only consider that it is not possible to tell from SIV/HIV-2 genetic sequences whether they were derived from an infected sooty mangabey or an infected human. We can add to this the fact that the range of sooty mangabeys in West Africa overlaps the epicenter of the HIV-2 epidemic (25,26). Moreover, several cases of apparent SIV infection of laboratory workers are now under investigation (27,28).

The close genetic relation of human and simian immunodeficiency viruses is apparent from the phylogenetic tree shown in Fig. 1. For the sake of clarity, only a very small number of known HIV and SIV sequences have been included in this tree analysis, which is based on relatively conservative homologous *pol* gene sequences (29). This tree is representative of the many sequences that have not been included,

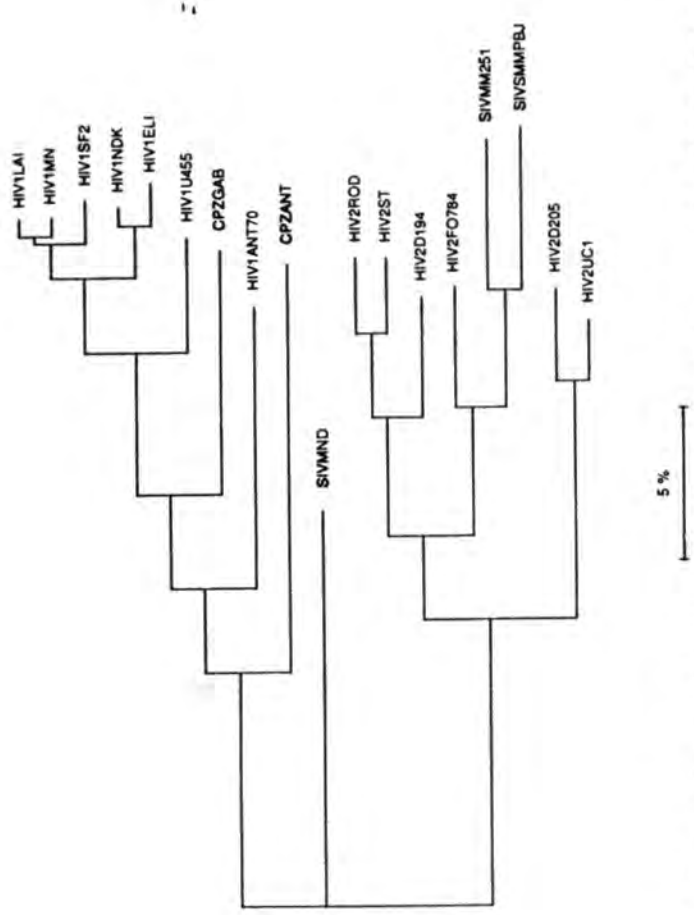


FIG. 1. Phylogenetic tree analysis comparing *pol* gene sequences of representative HIV-1, HIV-2, SIV, and CPZ samples. For the sake of clarity, only a very small number of known sequences have been included. Viral sequences derived from monkeys (SIV) and chimps (CPZ) have been highlighted. The HIV1ANT70 sequence may represent an animal viral sequence found on humans (see text). The tree was constructed by using the PAUP maximum parsimony algorithm. Of the 971 sites examined, 617 were varied. The lengths of the horizontal branches are proportional to single-base changes; these distances can be read as minimum percentage differences using the scale bar. The tree was rooted at the midpoint of the greatest patristic branch. Bootstrap analysis supported the tree topology as did tree analyses of other coding sequences.

and it is congruent with trees over other major viral-coding sequences. To emphasize some of the close relationships that are being encountered by AIDS researchers, SIV-derived sequences are highlighted in Fig. 1. Note that an "HIV-2" sample, designated F0784 in Fig. 1, taken from an asymptomatic Liberian individual (30), clusters with SIVs in the computer analysis. This is one of the most compelling pieces of evidence in support of the suspected mangabey virus origin of HIV-2. The close linkage of HIV-1 sequences with viral sequences isolated from wild-caught chimpanzees is the best clue yet for the origin of HIV-1s (31,32). It is apparent from Fig. 1 that a Cameroonian "HIV-1" isolate, denoted ANT70 (33), could be interpreted as a chimpanzee virus found in a human.

A highly diverse reservoir of immunodeficiency viruses have been catalogued in African green monkeys and their relatives (15,34,35). Although the genetic sequences determined thus far for these species do not directly account for the emergence of the two known types of HIV, recent phylogenetic studies suggest that the mangabey viruses found in West African animals are descendants of *sabaeus* monkey viruses also found in West Africa (Hahn B, personal communication). At least 13 species of feral African primates are known to be carrying HIV-related viruses (15,34); hence, additional zoonotic infections are clearly possible.

The epidemiological and epizootic moments implicit in Fig. 1 are historical and evolutionary turning points, suggestive of a discontinuity in the evolutionary flow. When there has been a breakthrough of this sort, Gould argues (as restated by Allan) that "punctuated equilibrium asserts that the rate of evolution is accelerated during times of environmental upheaval, resulting in the formation of new species" (15). This theory, which has traditionally been understood as *adaptive radiation*, has been described by Gould and colleagues in exact mathematical terms (36); referring to this process as *bottom-heavy cladism*, they predict a burgeoning of genetic forms during the period of evolutionary adjustment to a new host (here, humans). There is, in fact, evidence for rapid speciation of both HIV and SIV in captive animals concurrent with the AIDS pandemic.

Rapid Speciation Indicates Disequilibrium

If one looks closely at the "leaves" of the phylogenetic tree shown in Fig. 1, one sees that there are at least three, distinct subtypes of HIV-2. It is not now known whether these subtypes arose through independent events of cross-species transmission, although that seems the likeliest possibility. Additional information about West African simian immunodeficiency viruses is needed to answer this question. The subtypes represented by HIV2ROD and HIV2D205 in Fig. 1 have an average difference of 23% in their coding sequences (37), and within each subtype distances of up to 18% are encountered (29). This degree of diversification indicates that the subtypes will probably evolve into distinct viral types, differing by at least 50% in their coding regions.

From newly determined HIV-1 *gag* and *env* gene sequences (29,38–40), five distinct sequence subtypes have now been identified in the major centers of the AIDS pandemic (Fig. 2;29). These HIV-1 subtypes appear to have diverged from one another as recently as 1960, that is to say, they may have proliferated in step with the pandemic. (The lookback calculations supporting this hypothesis are considered in the next section of this chapter.) Because the surveillance effort is still in its infancy, it is conceivable that additional prevalent viral sequence subtypes will be discovered in the next few years.

Many countries—Brazil, Thailand, Uganda, and India, to name a few—now have cocirculating strains of HIV-1. The fact that only a single sequence subtype of the five or more known subtypes has yet been reported in United States AIDS patients is the result, we suppose, of a very strong founder effect. Currently, no one subtype seems to be more successful or more virulent (but, see later section on variation and pathogenesis), nor do the subtypes appear to have host preferences. The measurements necessary for evaluation of such phenotypic viral polymorphism prove quite difficult, however, and subtle changes in transmissibility and the average incubation time for AIDS can be discerned only with data gathered over a period of many years. Although trafficking of extant HIV-1 subtypes will undoubtedly make a greater contribution than mutation to the virus's ecological variation in the coming decade, the future spectrum of HIV phenotypes, we will argue, will be shaped by mutation (including recombination) and selection.

Transmission Contributes to Viral Diversification

Gould's theory of *bottom-heavy cladism* attempts to account for temporal asymmetries in the speciation process over long time periods. Although the existence of the five HIV-1 sequence subtypes shown in Fig. 2 is consistent with it, given the short time HIVs and SIVs have been observed, the HIV sequence subtypes are not sufficient to confirm the theory. Even so, the relationships of viruses shown in Fig. 2 are suggestive of a condition of disequilibrium that cannot be simply explained by a high mutation rate. Influenza A virus, for example, mutates about as rapidly as HIV-1, but its pattern of evolution is entirely different from that represented in Figs. 1 and 2: Flu-A trees are decidedly slender, whereas HIV-1 trees are bushy (11,41), which reminds us of Coffin's dictum that "mutation and variation are not the same" (42). This difference stems, for the most part, from the relative lack of intraspecific competition among HIV variants infecting resting cells. It is as if the niche for HIV within humans is more tolerant than that for influenza A. Until a steady-state evolutionary process is attained, HIV-1 variation might be epidemic-driven in a way that flu-A variation is not: specifically, the rate of HIV variation could be a nonlinear

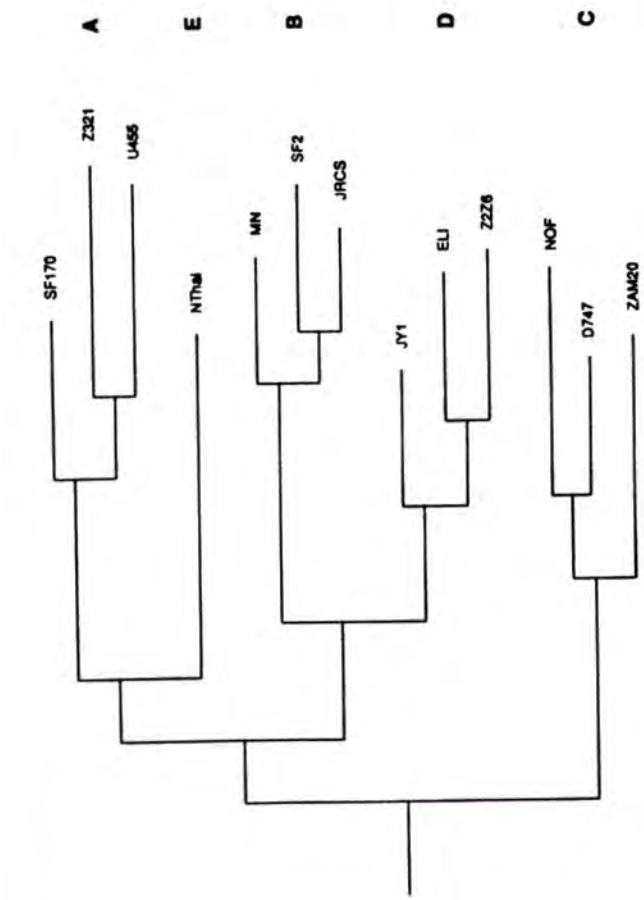


FIG. 2. Phylogenetic tree analysis comparing *env* gene sequences of diverse HIV-1 samples. The clades that are approximately equidistant have been tentatively recognized as sequence subtypes *A* through *E*. The tree was constructed as described for Fig. 1; it is virtually congruent with a phylogenetic tree based on *gag* gene sequences (29,38). At this still early time in the pandemic, subtype *A*, *C*, and *D* viruses are found preponderantly in Africa; subtype *B* viruses are found mostly in the United States and Europe; and subtype *E* viruses are characteristic of the outbreak in northern Thailand (but are also found in Central African Republic). The ANT70 sequence, included in Fig. 1, could represent a distinct HIV-1 subtype.

function of the annual AIDS incidence rate and the total number of infected individuals, irrespective of the hypothesized simian origin and any general concept of temporal asymmetry of viral adaptation. (The Luria–Delbruck equations, which describe a Darwinian process for rapidly mutating bacterial populations, turn out to be a useful starting point for exploration of this possibility.) The extent, if any, to which HIV-1 variation is epidemic-driven, should become apparent as a dilation of branch lengths in phylogenetic trees, perhaps by the year 2000.

Finally, the genetic proliferation of HIV in the coming decades can be considered from an ecological perspective. Ewald has argued that cultural characteristics will dictate, in very short times, the variability and virulence of a pathogen such as HIV (43,43a): "As frequencies of unprotected sexual contact with different partners increase, the [viral] benefits of increased durations of infectiousness should decline and virulence should increase." This prediction, which emphasizes variation through selection over mutability, can be corroborated or disproved in a few years, assuming that the critical timeframe of the HIV pandemic will be comparable with the outbreak and evolutionary resolution of the myxoma virus infection in Australian rabbits, a case thought to have some parallels with AIDS epidemicity. [Contrary to popular understanding, Ewald argues, the myxoma virus remains to this day significantly pathogenic in Australian rabbits.] Ewald proposes that "explaining the evolution of HIV thereore involves determination of how a benign pathogen [SIV] can evolve to high levels of virulence rather than how an extremely virulent pathogen [myxoma virus] can evolve to high levels of virulence" (43). The thrust of this argument is that human behavior—"sexual contact diversity" and intravenous drug use involving shared needles—will become an increasingly significant factor in the evolution and pathogenicity of HIV.

Whereas it has been argued by Temin that it is doubtful that HIV-1 will evolve into a worse pathogen, his argument is based on the assumption that the virus must simultaneously alter its infectiousness and its pathogenic potential, which he points out are "already optimized for spread and disease in the human population" (44). Temin goes on to suggest that HIV-1 may gradually evolve into a less pathogenic form. Ewald, on the other hand, theorizes that the pathogenic equilibrium could still shift upward as a result of human behaviors that provide a context for a decreased viral latency period. He also could imagine evolution toward a less pathogenic state, but only with the proviso of a more conservative human sexual behavior. Thus, the two outlooks differ in their assumptions about the major causal forces. HIV-2 disease, which already presents a greater clinical spectrum than HIV-1 disease while infecting far fewer individuals (see later section on variation and pathogenesis), provides an important context in which Ewald's causal hypothesis can be judged.

In this section we have reviewed the qualitative and semiquantitative evidence that supports the theory of a recent simian origin and rapid diversification of the HIVs. The probable implication of this viral history is that HIV will continue to be an atypically unstable and aggressive pathogen for some time. Before turning directly to questions about HIV pathogenicity, we will next review some relevant quantitative features of HIV molecular variation.

MOLECULAR FACTORS CONTRIBUTING TO HIV VARIATION

Rate of Variation

The theory of a recent simian origin and rapid speciation of HIV is supported by a determination of the average phenomenological rate of macroscopic variability. The possibility of determining such a quantity has been brought into question by Coffin, on the basis of widely observed intrapatient heterogeneity (42). The preceding discussion of HIV variation has emphasized global diversity, what can be called *macroscopic swarming*. Intrapatient heterogeneity, or *microscopic swarming*, is a second feature of diversification to be considered in any discussion of HIV evolution. In general, RNA viruses are highly mutable, but with RNA viruses such as foot-and-mouth-disease virus (FMDV), which infects cattle, and HIV, the variability is so great that there are populations of closely related viral genotypes within most infected individuals. These populations, or swarms, are termed *quasispecies* (4, 45,46; also see Chap. 10). Their relevance for progression to AIDS in HIV infections is a matter of some dispute (see section on variation and pathogenesis). Furthermore, it is unclear how selection comes into play when one individual infects another individual: since each transmission can be thought of as a recloning step that may or may not destabilize a quasispecies, Coffin has hypothesized that "cyclic patterns of variation" will arise and, therefore, that "there can be no such thing as a 'rate of variation'" (42). (Genetic bottleneck transmission, which depends on small innocula sizes, undoubtedly contributes to the indeterminancy emphasized by Coffin (47). While we agree with Coffin that empirical determinations of variation at a defined position in a coding sequence may prove intractable, this contention does not apply to macroscopic or global variation.

Molecular epidemiological investigations of mother–baby transmission, sexual transmission, and transmission from an infected dentist to his patients have revealed viral sequence similarities that were intermediate between those observed within a single individual and those observed between unlinked individuals (48–50). Furthermore, rate estimates of HIV-1 nucleotide substitution made by many investigators for many samples have been in surprisingly good agreement (reviewed in 11). On the basis of these findings, a continuous magnitude, representing an average phenomenological rate of macroscopic variability of HIV, may be determined. (Calculations based on this phenomenological rate presuppose a model akin to models of Brownian motion in which stochastic variables are handled as deterministic class-averages at the macroscopic level.)

To investigate this average quantity, we first determined frequencies of synonymous and nonsynonymous substitutions for a large number of HIV-1 *gag* and *env* gene sequences from a variety of infected individuals. Figure 3 shows a plot of the rates (52) of nonsynonymous (d_n) versus synonymous (d_s) nucleotide substitutions for 35 complete HIV-1 *env* gp120-coding sequences: d_s is the corrected frequency of synonymous substitutions that could have occurred for each of the 595 possible pairwise sequence relations that have been recorded in Figure 3. This value

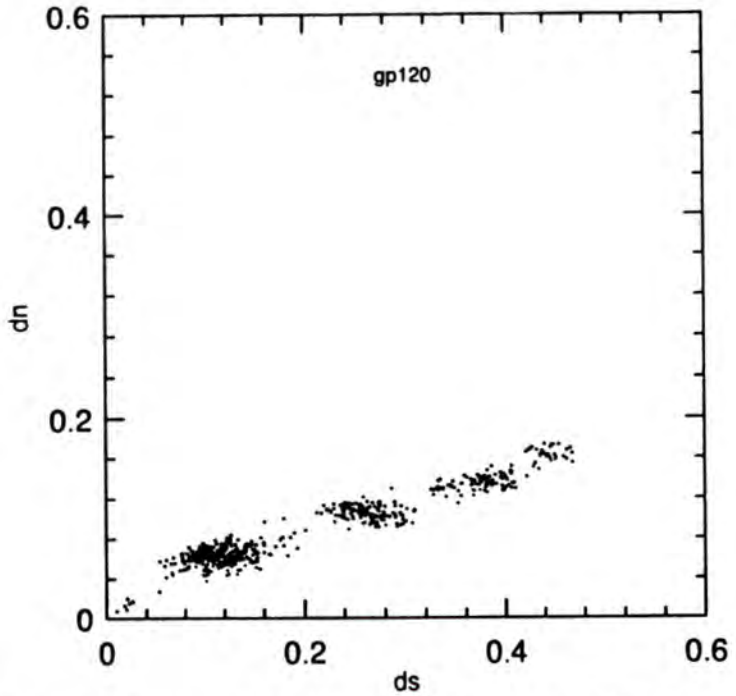


FIG. 3. Comparison of nonsynonymous (d_n) to synonymous (d_s) nucleotide substitutions in the HIV-1 *env* gp 120-coding sequences for 35 distinct samples. The Nei–Gojobori algorithm was employed (51). The slope, 0.4, is indicative of weak purifying selection. A similar analysis of HIV-1 *gag*-coding sequences (not shown) produces a slope of 0.2; for comparison, the influenza A hemagglutinin-coding sequence yields a slope of 0.3 (52). Only recently diverged *env* sequence relationships (d_s<0.5) are shown. The different sequence subtypes seen in Fig. 2 are responsible for the clustering effect. For relationships close to or greater than expected mutational saturation (d_s = 0.75), the slope becomes less than 0.4.

is, accordingly, a fair indicator of the mutational distance separating relatively close sequences, such as those represented in the *env* gene tree in Fig. 2 (d_s<0.5). More distant relations involving the CPZ and ANT70 sequences (included in Fig. 1) have synonymous substitution values of about 0.75, the expected saturation state, and are not shown in Fig. 3.

The *gag* and *env* substitution frequencies (d_s) represented in these studies, taken together with similar rates determined by several investigators (reviewed in 11), support the hypothesis that the five recognized HIV-1 sequence subtypes (see Fig. 2) diverged from one another as recently as 1960, a plausible date for the onset of the AIDS pandemic. The reasoning is as follows: the average intertypic nucleotide substitution rate for HIV-1 *env* divergence is about 1.1% per year (11; and see Fig. 3), and the average intertypic nucleotide distance, the Hamming distance, shown in Fig. 2 is about 32%, suggesting a lookback time of about 30 years. The *gag* rate is

just half the *env* rate (as is the ratio of d_n to d_s for *gag*-coding sequences) and the average intertypic distance is about 14% (29,38). With this same value for the *gag* fixation rate, Khan and colleagues calculated that the stump-tailed macaque virus apparently diverged from captive sooty mangabey and rhesus macaque viruses as recently as 1963, the approximate date of arrival of these animals into the colonies housing those other animals (23).

These estimates are consistent with the historical and seroepidemiologic records discussed in the earlier section on emergence of HIV. Hence, without evidence to the contrary, average rates of variation derived from large data sets suitably constituted from recently diverged viral sequences may be taken to have useful predictive value. However, the most important indicator for future HIV variation will be the nonsynonymous nucleotide substitution rate (d_n), which is extraordinarily high. This is apparent from Fig. 3 in which we see first, that a linear relation of d_n to d_s is preserved for values of d_s between 0.05 and 0.5, and second, that the slope in Fig. 3 (0.4) is relatively steep, greater, for example, than the equivalently determined slope for the influenza A hemagglutinin gene (0.3) (52). The ratio of d_n to d_s determined for 45 HIV-1 *gag*-coding sequences was approximately 0.2 (data not shown). The unusually high level of nonsynonymous substitution in HIV *env* sequences has been documented by several investigators, although with fewer sequences than were analyzed in Fig. 3 (9,53–56).

The ease with which amino acid replacements take place in HIVs is explained by their optimal mutation rate (57) and their unusual codon usage favoring nonconservative changes (11), but it is also a consequence of weak constraints. By one interpretation, the fact that synonymous substitutions predominate, supports Kimura's "neutral theory" (52). From another perspective, the relatively high ratio found in the HIV-1 *env* gene could be taken as evidence for "positive Darwinian selection" (58), or for "overdominant selection" (56,59). This latter possibility seems particularly plausible in the context of microscopic HIV evolution for which selection by an infected individual's immune response may favor change: extremely large ratios have been observed among microscopic swarms (48,54). Although we shall not attempt to weigh these alternatives in this chapter, the high ratios of nonsynonymous to synonymous substitution, which are indicative of weak purifying selection, have broad evolutionary implications.

Breadth of Variability

Although we have discussed only *gag* and *env* gene variability, HIV proteins, in general, are subject to high frequencies of amino acid replacement. There is a common misperception about the genomic distribution of variation observed in HIV-1s and HIV-2s; namely, that it is limited, for the most part, to the hypervariable regions of the *env* gene. Most of the HIV regulatory and accessory proteins, as well as select portions of the *gag* polyprotein, are as variable as the envelope protein (9,11,29). These findings are evident in the comparisons of average information

TABLE 1. *Protein information densities expressed as amino acid equivalents*

Viral sequences	*gag*	*pol*	*tat*	*rev*	*env*	*nef*
HIV-1/CIV (8)	0.77	0.86	0.65	0.58	0.60	0.64
HIV-2/SIV (7)	0.84	0.90	0.55	0.70	0.73	0.55
HIV-1/HIV-2/SIVs (15)	0.56	0.62	—	0.30	0.35	0.34

The above data have been previously reported and discussed in refs. 9 and 11. The information density expresses the average number of invariant and chemically conserved residues at each position of the protein molecule. For example, there are 0.65 amino acid equivalents, on the average, at each position in the HIV-1/CPZ *tat* protein set under study. Comparisons within each row involve the same set of complete genomic sequences: 8 HIV-1/CPZs, 7 HIV-2/SIVs, and 15 combined HIV/SIVs. The values will decrease as more sequences are considered; in this discussion, the focus is on the ratios of protein variation. Thus HIV-1 *env* protein is actually no more variable than the HIV-1 *rev* protein (0.60 versus 0.58).

densities, expressed as amino acid equivalents, of homologous HIV proteins. Table 1, for example, provides a summary of *tat*, *rev*, and *nef* protein variability in comparison with the *gag*, *pol*, and *env* proteins; these three regulatory proteins are known to play crucial roles in the viral life cycle (60). The HIV-1 *rev* protein, for instance, is as variable as the HIV-1 *env* protein according to the measurements summarized in Table 1: on the average, HIV-1 *rev* protein manifests 0.58 amino acid equivalents at each residue position, versus 0.6 amino acid equivalents at each position in the *env* protein. The HIV-2 *tat*, *rev*, and *nef* proteins, all are as variable as their correlative *env* protein (0.55, 0.70, and 0.55, versus 0.73). The *rev* and *nef* values for the combined HIVs (0.30 and 0.34) are close to or less than the value for the combined *env* (0.35) and are also close to typical values encountered for eukaryotic proteins having very distant evolutionary relationships (11). In the previous analysis (see Fig. 3), we noted that HIV-1 *gag* gene variation was typical, but that *env* gene variation was not; the comparative data in Table 1 confirms the notion of weak purifying selection and extends it to several other HIV proteins. To the extent that HIV variation is more than merely a property of select (hypervariable) regions of the lentiviral envelope, balanced polymorphism becomes a major factor in HIV speciation.

Other molecular dimensions of HIV variability that undoubtedly contribute to its rich evolutionary potential are shown by the wide diversity of promoter–enhancer configurations and envelope glycosylation profiles: HIVs and SIVs may possess zero, one, or two NFκB-binding sites in their long-terminal repeat (LTR) U3 regions (35,61,62) and may differ in their envelope modification at as many as 25 potential *N*-linked glycosylation sites (9,29,63,64).* Thus, the HIVs are both degenerate and complex. The combination of variability, degeneracy, existence as quasispecies, and complexity in HIVs allows many different lineages, and possibly

*The HIV-1 envelope gp120 molecule is one of the most heavily glycosylated proteins in nature; the hypervariability of the *env* gene appears to be strongly tied to mutational changes in the potential modification sites, NXS and NXT, where X is almost any amino acid.

different evolutionary scenarios, under a variety of selective forces. Complementation within the quasispecies (65), heterologous transactivation (11,60,61), transduction (66, also see Chapter 10), pseudotyping (67), and "bottleneck transmission" (47) are some of the compounding determinants of variability in these highly plastic pathogens. In view of these considerations, it is highly likely, as we will argue in the following, that different phenotypic lineages of HIV will emerge.

HIV VARIATION AND PATHOGENESIS

In the first section of this paper we concluded with some certainty that the human AIDS-causing viruses have had a recent simian origin. From several points of view, this breakthrough was interpreted as a portent of viral genetic instability for some time to come. In the second section of the paper, some of the quantitative conditions of this instability were examined. We saw that the genetic volatility can rapidly scramble parts of the HIV *env* protein (as a result of a high ratio of nonsynonymous to synonymous substitutions; see Fig. 3) and that it involves the viral regulatory and accessory proteins as much as the viral coat (see Table 1). Accordingly, we can expect to see continued HIV speciation. What effect, if any, can we expect this to have on the global pathogenic potential of these viruses?

Two Causal Views of HIV Pathogenesis

To attempt to summarize all of the current hypotheses about HIV pathogenesis is beyond our scope. (For a recent review of HIV immunopathogenesis, see ref. 68.) We shall limit our discussion to two classes of hypotheses that are most instructive in relation to the questions we have posed. One category of explanations for HIV pathogenesis emphasizes genetic variation in only general terms; it argues that HIV genetic variation needs to be only pervasive and nonspecific to produce pathogenic effects. The second class of hypotheses, supported by a now sizeable body of evidence, points to the occurrence of specific genetic determinants of the pathogenic phenotype. With the first class of hypotheses, the emphasis is on immunogenicity as the pathological force. The second class turns its focus toward direct cytolytic properties traceable to the HIV or SIV genotype. These two classes of explanations for HIV pathogenesis are not mutually exclusive: cell killing can be both direct and indirect.

Several examples of the first class of hypotheses may be mentioned. It has been proposed that viral variation could play a supportive role in a process of clonal antibody dominance: dominant B- and T-cell clones that have arisen in response to an initially homogeneous viral antigen ("original antigenic sin") might be stimulated by subsequent viral variants, irrespective of their precise makeup, resulting in an antibody "repertoire freeze" (69). A second possibility here is that of a "superantigen effect" (70). In a third example that involves viral variation only in a broad

sense, HIV diversification is thought sufficient to exhaust the immune system: an essential asymmetry between antigen and antibody stems from the fact that all viral variants can threaten cells through interaction with the CD4 receptor, but each antibody clone is type-specific for a single variant (71,72). Many other mechanisms that fall under this category have been put forward, among which might be mentioned the hypothesis that AIDS is essentially an autoimmune disease (73). Common to all of these various theories is the concept that variation is significant only in a general way. As Haseltine puts it, "transmission, pathogenesis and incubation of the disease is similar regardless of the infecting strain" (74).

The foremost example of specific genetic determination of atypical AIDS pathogenesis is an acutely lethal SIV variant that arose after a single passage in a macaque. The parent virus in this case was benign in its natural host, the mangabey, but through mutating to the highly lethal form in the macaque—able to kill macaques in weeks rather than years—it acquired aggressive pathogenic properties when placed back into mangabeys (75,76). A less dramatic, but no less informative, example is encountered in the comparison of two HIV-2 variants found to be infectious in macaques. The HIV-2 variant with enhanced cytopathogenicity and replicative rate has significantly fewer envelope modification sites (the protein sites mentioned earlier) than the second, less cytopathic variant (77). Cases have been observed in which naturally occurring attenuated forms have yielded cytopathic variants (78) and, conversely, highly cytopathic forms have generated nonpathogenic viral progeny (79,80). We are reminded that AIDS neuropathogenesis is a distinct evolutionary concern when we learn about HIV-1 variants with different properties of gene expression in the mouse brain (81).

Evidence For Different Pathogenic Phenotypes of HIV

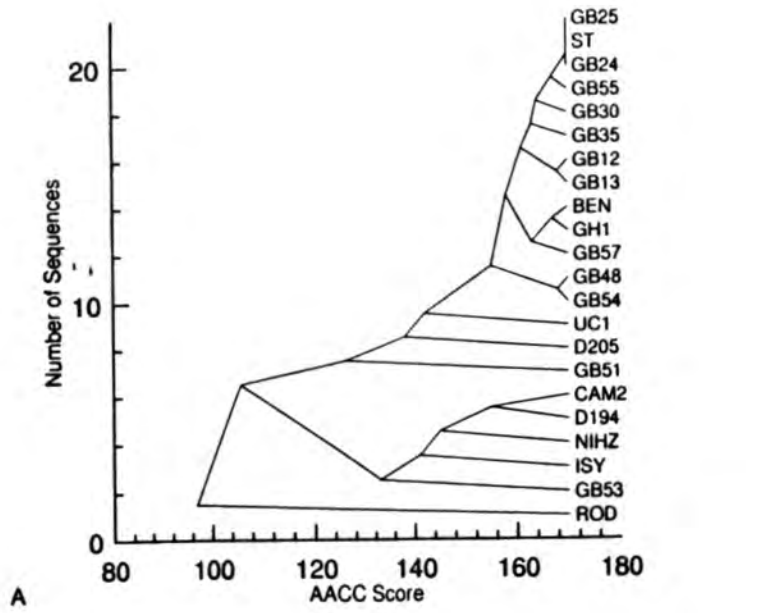
Although the strongest argument for differing viral phenotypes is derived mostly from animal studies, evidence exists for different pathogenic phenotypes of HIV in the global setting. Early in the epidemic, a Swedish research team noted that persons harboring viruses that could be characterized as having "rapid-high" phenotypes in vitro were closer to clinical AIDS than persons whose viruses were "slow-low" in culture (82). Further investigation of these two HIV-1 states—one predominant in early infection, the other predominant in late infection—has linked them, first, to the capacity for not inducing syncytia (NSI) versus inducing syncytia (SI) in vitro and, second, to abilty to grow on macrophages (macrophage-tropic or monocytotropic) versus ability to grow on transformed T-cell lines (T-cell–tropic) (83–91). Thus, generally speaking, progression to HIV-1 AIDS has been more rapid in persons harboring the SI viral phenotype that are said to be T-cell-tropic (83,84). The correlation is not without exceptions—for example, an extremely cytopathic virus that is also macrophage-tropic has been reported (92); however, this has been the finding of many investigators.

With few qualifications, the functionally critical determinant of the SI phenotype has been traced to a region of the viral envelope known as the V3 loop or principal neutralizing determinant (PND), a 35-amino acid peptide located in the third hypervariable domain of the envelope protein. No single region of the HIV genome dictates the variability underlying HIV phenotypic diversity, however, the V3 loop has received the greatest attention because (a) it serves as an epitope for potent immunological responses (93,94); (b) it affects fusion and viral virulence (85); and (c) it is implicated as a primary determinant of HIV-1 cell tropism (86,87). A single amino acid substitution in this variable protein can have phenotypic ramifications (88,89). Positively charged amino acid residues, most often arginines, found in the flanking regions of the V3 loop have been associated with the SI, or the more cytopathic, phenotype (90).

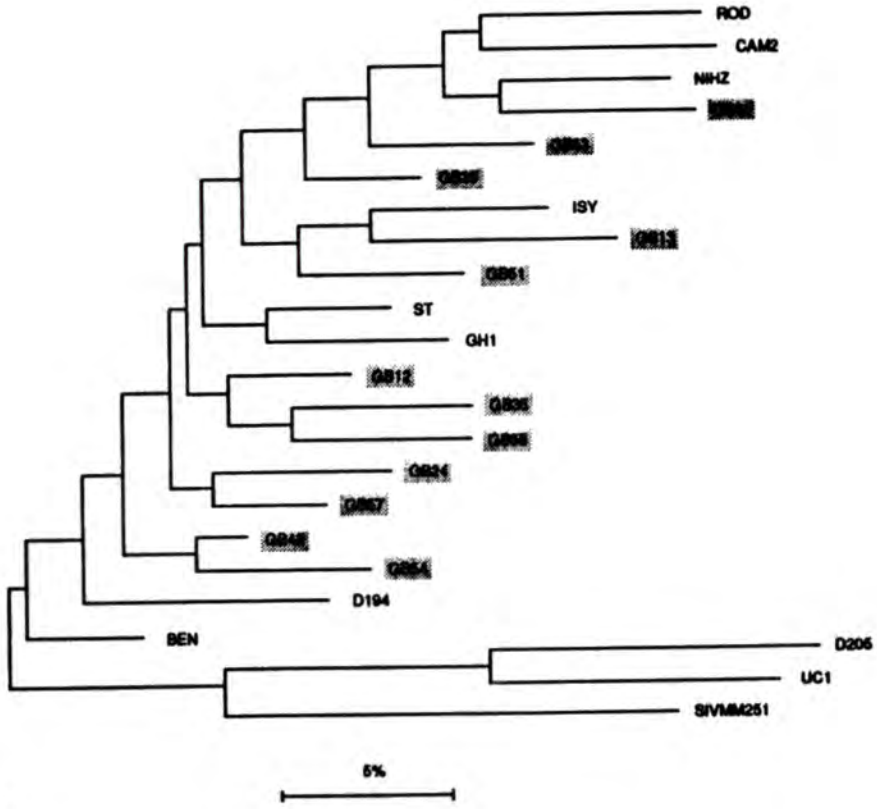
The correlation of V3 peptide-coding sequences with pathogenesis has almost entirely been focused upon HIV-1 variants, as exemplified by the hundreds of HIV-1 V3 sequences deposited in the HIV sequence database (29). Samples of HIV-1 subtype D viruses, characteristic of many infections in Central Africa (see first section and Fig. 2), have shown a higher frequency of the SI phenotype than have subtype B viruses, which are typical of HIV-1 infections in the United States. However, because the spectrums of opportunistic diseases are very different in the United States and Africa, and because viral sampling has not been systematically controlled, it is not possible to conclude simply that the one viral subtype is intrinsically more pathogenic than the other viral subtype.

Even so, *phenetic* analyses of V3 loop protein sequences, which emphasize structural similarities, irrespective of the evolutionary pathway, suggest that uncannily similar V3 loop proteins can be encountered in otherwise dissimilar nucleotide sequence subtypes (95). This manifestation of evolutionary parallelism can be viewed as sequence convergence, although we prefer to interpret it as simply lack of divergence: certain V3 sequence lineages, suggestive of the monocytotropic forms of HIV-1, appear to be more stable than other V3 lineages, especially those of the T-cell–tropic variety. Thus, some Ugandan V3 loop protein sequences of one *cladistically* determined sequence subtype are found to tightly cluster with United States V3 loop sequences of another sequence subtype (95,96).

Although they have not been studied with the same intensity, HIV-2 V3 loop peptide sequences yield an intriguing pattern that has a suggestive correlation with the health status of the person from which the viral sequence was derived. As shown in Fig. 4, diverse HIV-2 viral sequences evaluated from a cladistic point of view (see Fig. 4B) can, nevertheless, share common V3 loop proteins evaluated from a phenetic point of view (see Fig. 4A). Correlations of HIV-2 V3 protein sequences with SI and non-SI phenotypes have not been attempted; however, the majority of infected persons with divergent forms of the V3 protein shown in Fig. 4A were symptomatic, whereas the persons with the relatively stable form of V3 were mostly asymptomatic (11 of the 13 tightly clustered V3 sequences in Fig. 4A were taken from asymptomatic blood donors in Guinea Bissau) (97). As we have previously noted, positively charged amino acids in the flanks of HIV-1 V3 loops have been associated with the more cytopathic phenotype (90); therefore, it is noteworthy to



A



B

find an abundance of positive charges in the homologous regions of the divergent HIV-2 V3 loop sequences taken from the symptomatic individuals represented in the Fig. 4 analyses.

We think that there may be parallels between the apparently stable HIV-2 V3 lineage and the previously mentioned HIV-1 V3 loops that are similar at the protein level, despite being distant by cladistic analysis. Neither form possesses the positively charged flanking residues characteristic of more cytopathic HIV-1s. Moreover, highly conserved V3 loop protein sequences are encountered in SIVagm sequences and in two chimpanzee viral sequences, none of which appear to be pathogenic in their wild-caught hosts.

We tentatively conclude from these findings that viral diversification in and of itself is not a condition for pathogenicity. Diversification of the HIV V3-coding sequence appears to be a major determinant of, or a major effect deriving from, HIV pathogenicity. However, V3 variation cannot be the sole marker factor for progression to AIDS. Macaques infected with molecularly cloned SIV can progress to disease without any apparent dependence on the V3 loop (98,99). Although the focus of this discussion has been on molecular evolutionary parallelism, a large part of the story is the conspicuous lack of genetic stability, the diversification of the V3 peptide sequence in both HIV-1- and HIV-2-infected individuals who typically possess T-cell–tropic viruses and are symptomatic. Genetic flux, not conservation, is the predominant observation at this time.

CONCLUDING REMARKS

The clinical spectrum of HIV-2 disease in the world is now wider than the spectrum for HIV-1 disease. Some of the viral samples represented in the analyses of Fig. 4 were taken from persons with fullblown AIDS; others were taken from persons with only neurological impairments; still others represent what may prove to be benign forms of HIV-2. Although HIV-2 affects far fewer people, it sets the minimal boundary conditions, so to speak, for the HIV-1 pathogenic potential. The recent observation that a group of long-term survivors of HIV-1 infection harbor closely related viruses derived from the same individual is especially significant (100). The possibility of emergence of less pathogenic forms of HIV-1 (Temin's expectation, discussed in the earlier section on emergence of HIV) is prefigured by the less pathogenic forms of HIV-2 in Senegal, for example (101). What remains to

FIG. 4. A comparison of HIV-2 V3 loop protein sequence diversification, represented by a phenogram in **A**, to nucleotide sequence diversification, represented by a cladogram in **B**. Figure 4B was generated from the V3 region-coding sequences analogous to the construction of the trees shown in Figs. 1 and 2. Figure 4A was generated from V3 loop amino acid sequences using an "amino acid class covering" algorithm that clusters sequences on the basis of protein chemistry and that minimizes cladistic bias (95). Sequences at the top of **A**, are nearly identical (AACC score of 175) indicative of highly conserved V3 loop protein sequences primarily among samples from asymptomatic seropositive individuals from Guinea Bissau. This tight-clustering pattern is not observed in the phylogenetic reconstructions based on nucleotide sequences (highlighted in B). The comparatively divergent V3 loop sequences in **A**, joined by similarity scores less than 140, were derived from symptomatic patients (as discussed in the text).

be determined is the extent to which HIV-1s and HIV-2s will take more, not less, pathogenic courses in some societies with high rates of viral transmission (Ewald's expectation, as discussed in the section on emergence of HIV).

One of the most frequently asked questions concerns the route of transmission of HIVs: How likely is it that HIV could be transmitted through airborne droplets? There is no simple answer to this question. Some nonprimate lentiviruses are thought by some to be transmitted through sputum and the air, and one, the equine infectious anemia virus, is transmitted by flies. Although there is no reason to think that the modes of HIV transmission will change in the coming decades, what gets transmitted could become a different question. Wain-Hobson raises the disturbing possibility that HIV gene capture, and attendant oncogenesis, becomes increasingly possible as the AIDS incidence rate rises (see Chap. 10).

This Heraclitean picture is not complete without some mention of the profound ecological disturbances caused by HIV. Of these, the resurgence of tuberculosis, the leading cause of death in the world, is clearly the most serious (17,102,103). Although there is controversy about the extent to which emerging tuberculosis, and the increase in the number of strains resistant to multiple antibiotics, is due to AIDS, no one to our knowledge denies the connection.

ACKNOWLEDGMENTS

The HIV Sequence Database and Analysis Project, which has been the springboard for much of this inquiry, is funded by the AIDS Division of the National Institute of Allergy and Infectious Diseases through an interagency agreement with the US Department of Energy. Kersti MacInnes, Rhoshel Lenroot, April Gifford, and Miranda McEvilly assisted with the alignments and analyses. We heartily thank Lynda Myers for her editorial help. The CPZant and HIV1ANT70 sequences, analyzed in Fig. 1 were kindly provided before publication by Marleen Vanden Haesevelde and Innogenetics, Belgium, LAUR 93–475.

REFERENCES

1. Levine AJ. *Viruses*. New York: Scientific American Library; 1992: Chap. 7.
2. Grmek MD. *History of AIDS: emergence and origin of a modern pandemic*. Princeton: Princeton University Press; 1990. Maulitz RC, Duffin J, translators.
3. Diamond J. The mysterious origin of AIDS. *Nat Hist* 1992;9:24–29.
4. Eigen M. The AIDS debate. *Naturwissenschaften* 1989;76:341–350.
5. Eigen M, Nieselt-Struve K. How old is the immunodeficiency virus? *AIDS* 1990;4(suppl 1):S85–S93.
6. Temin HM. Evolution of retroviruses and other retrotranscripts. In: Bolognesi D, ed. *Human retroviruses, cancer and AIDS: approach to prevention and therapy*. New York: Alan R Liss; 1988:1–28.
7. Gojobori T, Moriyama EN, Ina Y, Ikeo K, Miura T, Tsujimoto H, Hayami M, Yokoyama S. Evolutionary origin of human and simian immunodeficiency viruses. *Proc Natl Acad Sci USA* 1990;87:4108–4111.

8. Sharp PM, Li W-H. Understanding the origins of AIDS viruses. *Nature* 1988;336:315.
9. Myers G, MacInnes KA, Korber B. The emergence of simian/human immunodeficiency viruses. *AIDS Res Hum Retroviruses* 1992;8:373–386.
10. Myers G, MacInnes K, Myers L. Phylogenetic moments in the AIDS epidemic. In: Morse SS, ed. *Emerging viruses*. Oxford: Oxford University Press; 1993:chap 12.
11. Myers G, Pavlakis GN. Evolutionary potential of complex retroviruses. In: Levy JA, ed. *The retroviruses*, vol 1. New York: Plenum Press; 1992:51–104.
12. Coffin JM. Structure and classification of retroviruses. In: Levy JA, ed. *The retroviruses*, vol. 1. New York: Plenum Press;1992:19–50.
13. Olmstead RA, Langley R, Roelke ME, Goeken RM, Adger-Johnson D, Goff JP, Albert JP, Packer C, Laurenson MK, Caro TM, Scheepers L, Wildt DE, Bush M, Martenson JS, O'Brien SJ. Worldwide prevalence of lentivirus infection in wild feline species: epidemiologic and phylogenetic aspects. *J Virol* 1992;66:6008–6018.
14. Querat G, Audoly G, Sonigo P, Vigne R. Nucleotide sequence analysis of SA-OMVV, a visna-related ovine lentivirus: phylogenetic history of lentiviruses. *Virology* 1990;175:434–447.
15. Allan JS. Viral evolution and AIDS. *J NIH Res* 1992;4:51–54.
16. Grmek MD. *History of AIDS: emergence and origin of a modern pandemic*. Princeton: Princeton University Press; 1990:128. Maulitz RC, Duffin J, translators.
17. Lederberg J, Shope RE, Oaks SC Jr., eds. *Emerging infections. Microbial threats to health in the United States*. Washington, DC: National Academy Press; 1992.
18. Merson MH. Slowing the spread of HIV: agenda for the 1990s. *Science* 1993;260:1266–1268.
19. Pfutzner A, Dietrich U, von Eichrel Un, von Briesen H, Brede HD, Maniar JK, Rubsamen-Waigmann H. HIV-1 and HIV-2 infections in a high-risk population in Bombay, India: evidence for the spread of HIV-2 and presence of a divergent HIV-1 subtype. *J AIDS* 1992;5:972–977.
20. Estebanez P, Sarasqueta C, Rua-Figueroa M, et al. Absence of HIV-2 in Spanish groups at risk for HIV-1 infection. *AIDS Res Hum Retroviruses* 1992;8:423–425.
21. Pieniazek D, Peralta JM, Ferreira JA, et al. Identification of mixed HIV-1/HIV-2 infection in Brazil by polymerase chain reaction. *AIDS* 1991;5:1293–1299.
22. Zekeng L, Salla R, Kaptue L, et al. HIV-2 infection in Cameroon: no evidence of indigenous cases. *J AIDS* 1992;5:319–320.
23. Khan AS, Galvin TA, Lowenstein LJ, Jennings MB, Gardner MB, Buckler CE. A highly divergent simian immunodeficiency virus (SIV-stm) recovered from stored stump-tailed macaque tissues. *J Virol* 1991;65:7061–7065.
24. Novembre FJ, Hirsch VM, McClure HM, Fultz PN, Johnson PR. SIV from stump-tailed macaques: molecular characterization of a highly transmissible primate lentivirus. *Virology* 1992;186:783–787.
25. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIV) closely related to HIV-2. *Nature* 1989;339:389–391.
26. Marx PA, Li Y, Lerche NW, Sutjipto S, Gettie A, Yee JA, Brotman BH, Prince AM, Hanson A, Webster RG, Desrosiers RC. Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a West African pet sooty mangabey. *J Virol* 1991;65:4480–4485.
27. Centers for Disease Control. Seroconversion to simian immunodeficiency virus in two lab workers. *MMWR* 1992;41:678–681.
28. Centers for Disease Control. Anonymous survey for simian immunodeficiency virus (SIV) seropositivity in SIV-laboratory researchers—United States, 1992. *MMWR* 1992;41:814–815.
29. Myers G, Korber B, Smith RK, Berzofsky J, Pavlakis G, eds. *Human retroviruses and AIDS 1992*. Los Alamos: Theoretical Biology and Biophysics, Los Alamos National Laboratory; 1992.
30. Gao F, Yue L, White AT, Pappas PG, Barchue J, Hanson AP, Greene BM, Sharp PM, Shaw GM, Hahn BH. Human infection by genetically diverse SIV_{SM}-related $HIV_{?}2$ in West Africa. *Nature* 1992;358:495–499.
31. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S. Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* 1990;345:356–359.
32. Peeters M, Fransen K, Delaporte E, Van den Haesevelde M, Gershey-Damet G-M, Kestens L, van der Groen G, Piot P. Isolation and characterization of a new chimpanzee lentivirus: simian immunodeficiency virus isolate CPZ-ANT from a wild-captured chimpanzee. *AIDS* 1992;6:447–451.
33. DeLeys R, Vanderborght B, Haesevelde MV, Heyndrick L, van Geel A, Wauters C, Bernaerts R, Saman E, Nijs P, Willems B, Taelman H, van der Groen G, Piot P, Tersmette T, Huisman JG, van Heuverswyn H. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of West-Central African origin. *J Virol* 1990;64:1207–1216.

34. Johnson PR, Myers G, Hirsch VM. Genetic diversity and phylogeny of nonhuman primate lentiviruses. In: Koff W, et al, eds. Annual review of AIDS research, vol 1. New York: Marcel Dekker;1991:47–62.
35. Hirsch VM, Dapolito GA, Goldstein S, McClure H, Emau P, Fultz PN, Isahakia M, Lenroot R, Myers G, and Johnson PR. A distinct African lentivirus from Sykes' monkey. *J Virol* 1993;67: 1517.
36. Gould SJ, Gilensky NL, German RZ. Asymmetry of lineages and the direction of evolutionary time. *Science* 1987;236:1437–1441.
37. Dietrich U, Adamski M, Kreutz R, Seipp A, Kuhnel H, Rubsamen-Waigmann H. A highly divergent HIV-2-related isolate. *Nature* 1989;342:948–950.
38. Louwagie J, McCutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy GA, van der Groen G, Fransen K, Guy-Dermat G-M, Deleys R, Burke DS. Comparison of *gag* genes from seventy international HIV-1 isolates provides evidence for multiple genetic subtypes. *AIDS* 1993;7:769–780.
39. Dietrich U, Grez M, von Briesen H, Panhans B, Geibendorfer M, Kuhnel H, Maniar J, Mahambre G, Becker WB, Becker MLB, Rubsamen-Waigmann H. HIV-1 strains from India are highly divergent from prototypic African and US/European strains but are linked to a South African isolate. *J AIDS* 1992;7:23–27.
40. Ou C-Y, Takebe Y, Luo C-C, Kalish M, Auwanit W, Bandea C, de la Torre N, Moore JL, Schochetman G, Yamazaki S, Gayle HD, Young NL, Weniger BG. Wide distribution of two subtypes of HIV-1 in Thailand. *AIDS Res Hum Retroviruses* 1992;8:1471–1472.
41. Buonagurio DA, Nakada S, Parvin JD, Krystal M, Palese P, Fitch WM. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. *Science* 1986;232: 980–982.
42. Coffin JM. Genetic diversity and evolution of retroviruses. *Curr Top Microbiol Immunol* 1992; 176:143–164.
43. Ewald PW. Transmission modes and the evolution of virulence. *Hum Nat* 1991;2:1–30.
43a. Ewald PW. The evolution of virulence. *Sci Am* 1993;268:86–93.
44. Temin HM. Is HIV unique or merely different? *J AIDS* 1989;2:1–9.
45. Steinhauer DA, Holland JJ. Rapid evolution of RNA viruses. *Annu Rev Microbiol* 1987;41:409–433.
46. Wain-Hobson S. HIV-1 quasispecies in vivo and ex vivo. *Curr Top Microbiol Immunol* 1992; 176:181–193.
47. Clarke DK, Duarte EA, Moya A, et al. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J Virol* 1993;67:222–228.
48. Wolinsky SM, Wike CM, Korber B, Hutto C, Parks W, Rosenblum L, Kuntsman K, Furtado M, Munoz J. Selective transmission of human immunodeficiency virus type I (HIV-1) variants from mothers to infants. *Science* 1992;255:1134–1137.
49. Burger H, Weiser B, Flaherty K, Gulla J, Nguyen P-N, Gibbs RA. Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity among close contacts. *Proc Natl Acad Sci USA* 1991;88:11236–11240.
50. Korber B, Myers G. Signature pattern analysis: a method for assessing viral sequence relatedness. *AIDS Res Hum Retroviruses* 1992;8:1549–1560.
51. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 1986;5:418–426.
52. Gojobori T, Moriyama EN, Kimura M. Molecular clock of viral evolution, and the neutral theory. *Proc Natl Acad Sci USA* 1990;87:10015–10018.
53. Leigh-Brown A, Monaghan P. Evolution of the structural proteins of human immunodeficiency virus: selective constraints on nucleotide substitution. *AIDS Res Hum Retroviruses* 1988;4:399–407.
54. Simmonds P, Balfe P, Ludlam CA, Bishop JO, Leigh-Brown AJ. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. *J Virol* 1990;64:5840–5850.
55. Burns DPW, Desrosiers RC. Selection of genetic variants of SIV in persistently infected rhesus monkeys. *J Virol* 1991;65:1843–1854.
56. Shpaer EG, Mullins JI. Rates of amino acid change in the envelope protein correlate with pathogenicity of primate lentiviruses. *J Mol Evol* 1993;37:57–65.
57. Nowak M. HIV mutation rate. *Nature* 1990;347:522.

58. Fitch WM, Leiter JME, Li X, Palese P. Positive Darwinian evolution in human influenza A viruses. *Proc Natl Acad Sci USA* 1991;88:4270–4274.
59. Hughes AL, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA* 1989;86:958–962.
60. Cullen BR. Human immunodeficiency virus as a prototypic complex retrovirus. *J Virol* 1991; 65:1053–1056.
61. Gaynor R. Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* 1992;6:347–363.
62. Ross EK, Bucker-White AJ, Rabson AB, Englund G, Martin MA. Contribution of NF-κB and Sp1 binding motifs to the replicative capacity of human immunodeficiency virus type 1: distinct patterns of viral growth are determined by T-cell types. *J Virol* 1991;65:4350–4358.
63. Myers G, Lenroot L. HIV glycosylation: what does it portend? *AIDS Res Hum Retroviruses* 1992; 8:1459–1460.
64. Lee W-R, Syu W-J, Du B, Matsuda M, Tan S, Wolf A, Essex M, Lee T-H. Nonrandom distribution of gp 120*N*-linked glycosylation sites important for infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci USA* 1992;89:2213–2217.
65. Li Y, Hui H, Burgess CJ, Price RW, Sharp PM, Hahn BH, Shaw GM. Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation. *J Virol* 1992;66:6587–6600.
66. Swain A, Coffin JM. Mechanism of transduction by retroviruses. *Science* 1992;255:841–845.
67. Le Guern M, Levy JA. Human immunodeficiency virus (HIV) type 1 can superinfect HIV-2-infected cells: pseudotype virions produced with expanded cellular host range. *Proc Natl Acad Sci USA* 1992;89:363–367.
68. Levy JA. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* 1993;57:183–289.
69. Kohler H, Goudsmit J, Nara P. Clonal dominance: cause for a limited and failing immune response to HIV-1 infection and vaccination. *J AIDS* 1992;5:1158–1168.
70. Coffin JM. Superantigens and endogenous retroviruses: a confluence of puzzles. *Science* 1992;255: 411–413.
71. Nowak MA, Anderson RM, McLean AR, Wolfs TFW, Goudsmit J, May RM. Antigenic diversity thresholds and the development of AIDS. *Science* 1991;254:963–966.
72. Nowak MA. Variability of HIV infections. *J Theor Biol* 1992;155:1–20.
73. Dalgleish AG, Collizi V. Role of major histocompatibility complex recognition in the protection and immunopathogenesis of AIDS. *AIDS* 1992;6:523–525.
74. Haseltine WA. Human immunodeficiency virus (HIV) gene expression and function. In: Gallo RC, Jay G, eds. *The human retroviruses*. San Diego: Academic Press; 1991:69–106.
75. Dewhurst S, Embretson JE, Anderson DC, Mullins JI, Fultz PN. Sequence analysis and acute pathogenicity of molecularly cloned SIV-smm-pbj14. *Nature* 1990;345:636–640.
76. Martin MA. Fast-acting slow viruses. *Nature* 1990;345:572–573.
77. Misher L, Schmidt A, Steele J, Morton W, Hu S-L. In vivo passage of HIV-2 EHO in *Macaca nemestrina* results in persistent infection, disease progression and altered cellular tropism. Presented at the 10th Annual Symposium on Nonhuman Primate Models for AIDS, San Juan, Puerto Rico, 17–20 November, 1992.
78. Hoxie JA, Brass LF, Pletcher CH, Haggarty BS, Hahn BH. Cytopathic variants of an attenuated isolate of human immunodeficiency virus type 2 exhibit increased affinity for CD4. *J Virol* 1991; 65:5096–5101.
79. Courgnaud V, Laure F, Fultz PN, Montagnier L, Brechot C, Sonigo P. Genetic differences accounting for evolution and pathogenicity of simian immunodeficiency virus from a sooty mangabey monkey after cross-species transmission to a pig-tailed macaque. *J Virol* 1992;66:414–419.
80. Cheng-Mayer C, Shioda T, Levy JA. Host range, replicative, and cytopathic properties of human immunodeficiency virus type 1 are determined by very few amino acid changes in *tat* and gp120. *J Virol* 1991;65:6931–6941.
81. Vasalloy JR, Buzy JM, Zink MC, Clements JE. Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. *Science* 1992;258:1804–1808.
82. Aspo B, Morfeldt-Månson L, Albert J, Biberfeld G, Karlsson A, Lidman K, Fenyo EM. Replicative capacity of human immunodeficiency virus from patients with varying severity of HIV infection. *Lancet* 1986;2:660–662.
83. Schuitemaker H, Koot M, Kootstra NA, Wouter Dercksen M, de Goede REY, van Steenwijk RP,

Lange JMA, Eeftink Schattenkerk JKM, Miedema F, Tersmette M. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus populations. *J Virol* 1992;66: 1354–1360.
84. Boucher CAB, Lange JMA, Miedema FF, Weverling GJ, Koot M, Mulder JW, Goudsmit J, Kellam P, Larder BA, Tersmette M. HIV-1 biological phenotype and the development of zidovudine resistance in relation to disease progression in asymptomatic individuals during treatment. *AIDS* 1992;6:1259–1264.
85. Page KA, Stearns SM, Littman DR. Analysis of mutations in the V3 domain of gp160 that affect fusion and infectivity. *J Virol* 1992;66:524–533.
86. Westervelt P, Gendelman HE, Ratner L. Identification of a determinant within the human immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes. *Proc Natl Acad Sci USA* 1991;88:3097–3101.
87. Hwang SS, Boyle TJ, Lyerly HK, Cullen BR. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science* 1991;253:71–74.
88. Wolfs TFW, Zwart G, Bakker M, Valk M, Kuiken CL, Goudsmit J. Naturally occurring mutations within HIV-1 V3 genomic RNA lead to antigenic variation dependent on a single amino acid substitution. *Virology* 1991;185:195–205.
89. Takeuchi Y, Akutsu M, Murayama K, Shimuzu N, Hoshino H. Host range mutant of human immunodeficiency virus type 1: modification of cell tropism by a single point mutation at the neutralization epitope in the *env* gene. *J Virol* 1991;65:1710–1718.
90. Fouchier RAM, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema HG, Schuitemaker H. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol* 1992;66:3183–3187.
91. Tersmette M, Gruters RA, de Wolf F, de Goede REY, Lange JMA, Schellekens PTA, Goudsmit J, Huisman HG, Miedema F. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential isolates. *J Virol* 1989;63:2118–2125.
92. Collman R, Balliet JW, Gregory SA, Friedman H, Kolson DL, Nathanson N, Srinivasan A. An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type-1. *J Virol* 1992;66:7517–7521.
93. Javaherian K, Langlois AJ, LaRosa GJ, Profy AT, Bolognesi DP, Herlihy WC, Putney SD, Matthews TJ. Broadly neutralizing antibodies elicited by the hypervariable neutralizing determinant of HIV-1. *Science* 1990;250:1590–1593.
94. Emini EA, Schleif WA, Nunberg JH, et al. Prevention of HIV-1 infection of chimpanzees by gp120 V3 domain-specific monoclonal antibody. *Nature* 1992;355:728–732.
95. Korber B, Myers G. Contrasting HIV phylogenetic relationships and V3 loop protein similarities. In: Girard M, ed. *Proceedings of the Septieme Colloque Des Cent Gordes*. Paris, France; Merieux Foundation; 1993:151–158.
96. Albert J, Franzan L, Jansson M, Scarlatti G, Kataaha PK, Katabira E, Mubiro F, Rydaker M, Rossi P, Pettersson U, Wigzell H. Ugandan HIV-1 V3 loop sequences closely related to the US/European consensus. *Virology* 1992;190:674–681.
97. Boeri E, Giri A, Lillo F, Ferrari G, Varnier OE, Ferro A, Sabbatani S, Saxinger WC, Franchini G. In vivo genetic variabilitiy of the human immunodeficiency virus type 2 V3 region. *J Virol* 1992;66:4546–4550.
98. Overbaugh J, Rudensy LM, Papenhausen MD, et al. Variation in simian immunodeficiency virus *env* is confined to V1 and V4 during progression to simian AIDS. *J Virol* 1991;65:7025–7031.
99. Johnson PR, Hamm JE, Goldstein S, et al. The genetic fate of molecularly cloned simian immunodeficiency virus in experimentally infected macaques. *Virol* 1991;185:217–228.
100. Learmont J, Tindall B, Evans L, Cunningham A, Cunningham P, Wells J, Penny R, Kaldor J, Cooper DA. Long-term symptomless HIV-1 infection in recipients of blood products from a single donor. *Lancet* 1992;340:863–867.
101. Romieu I, Marlink R, Kanki P, M'Boup S, Essex M. HIV-2 link to AIDS in West Africa. *J AIDS* 1990;3:220–230.
102. Bloom BR, Murray CJL. Tuberculosis: commentary on a reemergent killer. *Science* 1992;257: 1055–1062.
103. Bermejo A, Veeken H, Berra A. Tuberculosis incidence in developing countries with high prevalence of HIV infeciton. *AIDS* 1992;6:1203–1206.

The Evolutionary Biology of Viruses, edited by Stephen S. Morse.
Raven Press, Ltd., New York © 1994.

11

Evolution of Genetic Exchange in RNA Viruses

Lin Chao

Department of Zoology, University of Maryland, College Park, Maryland 20742

Evolution provides the conceptual framework for understanding biological diversity. In this chapter, I will discuss, from an evolutionary viewpoint, the process of genetic exchange in RNA viruses. From that perspective, genetic exchange and sex are synonymous (1) and sex is surprisingly similar in RNA viruses and eukaryotes. In both types of organisms, the same hypotheses can account for the evolution of sex, and equivalent problems arise.

SEX IN DNA VIRUSES

In all viruses, sexual reproduction occurs when two or more viruses coinfect the same host cell, and hybrid progeny are produced through genetic exchange between the coinfecting parent viruses.

In DNA viruses, genetic exchange is prompted by *recombination*, which is defined here as the generation of a new nucleotide strand containing information from two or more parental strands. Recombination in DNA viruses, just as in eukaryotes, is by a mechanism of breakage-and-reunion (2,3). Following a physical break in the parental DNA molecules, pieces from different parental molecules can be joined to form a recombinant. Breakage-and-reunion does not require replication of the parental DNA (4).

SEX IN RNA VIRUSES

Mechanisms of Genetic Exchange

Although sex and recombination are universal in DNA viruses, they are not in RNA viruses. Despite intensive searches, recombinants have never been identified in RNA bacteriophages and several eukaryotic RNA viruses [e.g., vesicular stomatitis (VSV) and Newcastle disease virus of fowl] (5,6). The ability of detecting

To be published in
Emerging Viruses (S.S. Morse, ed.),
Oxford Univ. Press, 1993
pp. 120-137

12

Phylogenetic Moments in the AIDS Epidemic

GERALD MYERS, KERSTI MacINNES, AND LYNDA MYERS

INTRODUCTION

By 1983, several laboratories had identified and isolated the etiologic agent of AIDS, the retrovirus now denoted HIV, and by 1985 nucleotide sequences derived from those early isolates were reported. This molecular information was available in time to address a flurry of speculations and allegations concerning the sudden emergence of the AIDS virus. In 1986, two avirulent strains of herpes simplex virus were discovered to have generated a lethal recombinant in vitro (Javier et al., 1986), and this quickly touched off speculation that the AIDS virus may have been similarly generated in nature or in a testtube. Such speculations contributed to a worldwide climate of anxiety and created a potential for dissemination of false and misleading information. In light of this, we at the HIV Sequence Database and Analysis Project (an NIH-funded project at the Los Alamos National Laboratory), recognized the immediate need for an investigation of the molecular facts pertinent to the origin of AIDS.

Of the many ways to inquire into the evolutionary origins of an organism or pathogen, a comparative study of genetic sequences is the most incisive, since the sequences constituting the genome are unique for each species and, in some cases, for each individual in a species. The HIV genome, which consists of a single molecule, is approximately 9,700 nucleotides in length, a length more than adequate for precise identification of the virus. Thus in 1985, we set out to classify the newly determined nucleotide sequences of HIV by performing a computer search of all known genetic sequences that were stored in international databases. The sequences of HTLV-III and LAV, as they were called then, were compared to approximately six million bases of mammalian, plant, and viral sequences with the result that no similarity could be found. The search for close relatives was repeated again in 1986 when there were nine million bases in libraries and again in 1988 when there were over fifteen million bases. This inquiry has shown that if HIV arose in recent time as a recombinant of two or more

viruses, naturally or through human agency, it came from viruses that have not been characterized, that is from genetic sequences that are not in the international databases. A growing body of circumstantial evidence—outlined below—argues that HIV has not arisen in recent decades as a recombinant of any viruses, known or unknown.

An equally provocative early speculation about the origin of AIDS was that it arose from a monkey virus that had recently crossed species lines. Unfortunately, this possibility was brought forward more than once in terms that were unnecessarily offensive to Africans, on the one hand, and not helpful to AIDS researchers on the other hand. In what follows we will present what we believe to be a more informed hypothesis positing a simian origin of AIDS. Arguments for this hypothesis will entail tracking HIV and SIV nucleotide sequences into the past, but before we turn to that project, we will mention some features of the onset of the AIDS pandemic which are of central importance for any discussion of the origin of AIDS.

EPIDEMIOLOGICAL SNAPSHOTS

Early seroepidemiological and clinical data reveal two especially perplexing features of the onset of the AIDS pandemic:

First, all available evidence suggests that AIDS appeared suddenly and concurrently in the United States and Africa, the two major centers of the pandemic. AIDS was first recognized in the United States in the spring of 1981. The retrospective examination of sera collected in the late 1970s in association with hepatitis B studies in San Francisco, Los Angeles, and New York City, suggests that for the most part HIV-1 entered the U.S. population sometime in the mid-1970s (Jaffe et al., 1985; Stephens et al., 1986). Thus, for example, stored sera of nearly one thousand Los Angeles patients having either acute or chronic hepatitis were examined for HIV antibodies and the earliest sample in which antibodies were detected was from 1979 (DeCock et al., 1987). Although there is the intriguing report of a young man from St. Louis who presented AIDS-like symptoms in 1968, and whose tissues, stored upon death in 1969, have tested positive for antibodies, confirmation of HIV infection by PCR analysis and sequencing has not been possible up to now. A review of serosurveys associated with dengue fever in the Caribbean likewise found that the earliest evidence of HIV in Haiti comes from 1979 samples (Pape et al., 1987).

A similar and virtually concurrent epidemiological record is found in the other large center of the pandemic, Africa. While the seroepidemidemiologic and clinical data are limited, the available facts, summarized in Figure 12.1, also point to a relatively recent eruption of AIDS there (Piot et al., 1988).

In Africa, just as in St. Louis, there is evidence of an exceptionally early case of AIDS: a 1959 seropositive sample from Kinshasa that was, unfortunately, never tested for virus (Nahmias et al., 1986). However, sera sampled in 1970 that were taken from 805 healthy women in the same city revealed only two positives, while the 1980 sampling of a smaller group revealed fifteen (Desmyter et al., 1986). Sera from 659 persons examined in connection with an 1976 outbreak of Ebola fever in rural Zaire produced five HIV positives. Of these, the one confirmed positive (it is well known that false positives increase with time of storage of serum samples) was known to have recently lived in Kinshasa where she may have worked as a prostitute (Getchell et al., 1987). The HIV-1 virus isolated from that stored 1976 sample is the oldest molecular specimen that we have (Srinivasan et al., 1989).

From nearby Congo, sera of 340 pygmies were collected between 1975 and 1978. These people are in direct contact with monkeys, which they hunt and eat. None of the sera from these pygmies has been confirmed HIV positive (Rouzioux et al., 1986). And in South Africa, 3,500 samples were taken from the black miner population between 1970 and 1974 as part of a pneumococcal vaccine study. At most, two of these (both unconfirmed) were HIV positive. Yet by 1986, nearly four percent of the subpopulation of miners from Malawi were found to be positive (Sher et al., 1987). In urban Somalia to the north, there is no evidence of HIV prior to 1983 (Titti et al., 1986). And an extensive study of prostitutes and patients at STD clinics in Nairobi revealed low prevalence of HIV positives in 1980 that rose to 60% in some groups by 1985 (Piot et al., 1986). By 1991, the World Health Organization estimated that six million adults were HIV infected in Africa, with some countries showing seroprevalence rates of greater than 10% (Palca, 1991).

These serological studies all point to a sudden recent spread of HIV by travelers and prostitutes in central Africa that was undoubtedly facilitated by the road running from outside Kinshasa, Zaire to Mombasa, Kenya (Fig. 12.1; Conner and Kingman, 1988). Thus, AIDS appeared suddenly and concurrently in the United States and Africa, the two major centers of the pandemic.

The second feature of the AIDS epidemic that is especially intriguing is the fact that two major foci of AIDS are manifest in Africa: HIV-1 disease in central and southern Africa and HIV-2 disease in west Africa. Although there is a paucity of retrospective data pertaining to HIV-2 that would show how closely the onset of the HIV-2 epidemic coincided in time with that of HIV-1, available evidence suggests that the temporal onset of the two epidemics may have been separated by years but hardly by several decades. Through review of serological samples from an investigation of Lassa fever in Guinea Bissau, Fultz and coworkers concluded that the seroprevalence of HIV-2 in west Africa, which was 6% in 1986, could have been no higher than 1% in 1980 (Fultz et al., 1987). There was no evidence of HIV-1 in west Africa until recently when a case of dual HIV-1/HIV-2 infection was reported from Ivory Coast (Rayfield et al., 1988). Without evidence to the contrary, we may assume for the sake of inquiry that AIDS has been a

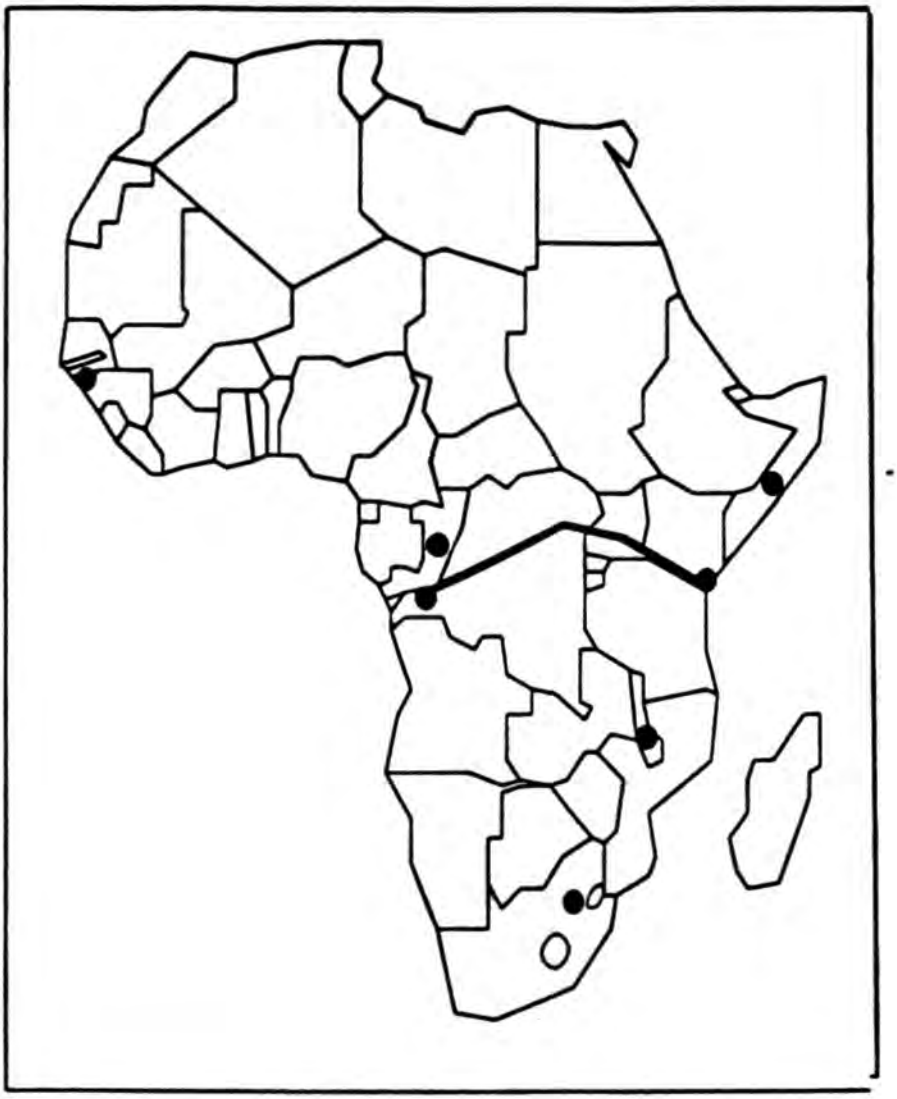


Figure 12.1 A computer-generated map of Africa showing sites for which retrospective serosurvey data, discussed in the text, is available. The main highway connecting Kinshasa, Zaire (in the west) to Mombasa, Kenya (in the east) is shown as pointed out by Conner and Kingman (1988).

"twofold epidemic," that HIV-1 disease and HIV-2 disease came forward at about the same time.

This brief epidemiological snapshot of the early years of the AIDS pandemic leaves us with two major questions: (1) What is implied by the suddenness and virtual simultaneity of onset of the epidemics in the United States and in Africa? (2) Under what presumption can the twofold (HIV-1 and HIV-2) epidemic in Africa be understood? We will turn to these questions now.

EVIDENCE FOR THE SIMIAN ORIGIN OF AIDS

In a succinct review entitled "The simian-human connection," Doolittle (1989) has summarized two recently popular hypotheses concerning the origin of HIV. The first is that HIV-2 (the second type of human AIDS-causing virus discussed above) derives from the mangabey virus. In support of this hypothesis is the fact that the nucleotide sequence of the mangabey virus proves to be closely homologous with HIV-2 sequences (Hirsch et al., 1989; Marx et al., 1991), and the fact that the natural habitat of the mangabey overlaps west Africa, the region of HIV-2 epidemic (Hirsch et al., 1989; Marx et al., 1991; Johnson et al., 1991). The second currently popular hypothesis presented by Doolittle, which we shall see is less likely, is that the other type of human AIDS-causing virus, HIV-1, recently arose from viruses expected to be found in an extensive population of SIV-positive green monkeys inhabiting sub-Saharan Africa (Hendry et al., 1986; Kanki et al., 1985; Johnson et al., 1991). How plausible are these hypotheses taken together? They require at least two distinct accidents of cross-species transmission over apparently the same time period.

Duesberg's challenge to the conclusion that HIV is the cause of AIDS raises precisely this problem: "It is highly improbable that within the last few years two viruses that are only 40% sequence-related would have arisen that could both cause the newly defined syndrome AIDS. ... [S]ince viruses, like cells, are the products of gradual evolution, the proposition that within a very short evolutionary time, two different viruses capable of causing AIDS would have evolved or crossed over from another species is highly improbable" (Duesberg, 1989).

While Duesberg's conclusion that HIV is not the cause of AIDS seems unwarranted for many reasons that we will not explore here, the logical improbability that he mentions must be faced. In order to do this, we shall make use of phylogenetic trees constructed from homologous gene sequences with the help of a computer. Figure 12.2 presents an example of a phylogenetic tree that has been built from representative *gag* gene sequences of the now known five types of primate immunodeficiency viruses, namely HIV-1 and the related SIV from a chimpanzee (IV); HIV-2 and its close relatives SIV-macaque and SIV-mangabey (I); SIV-AGM (African green monkey) (III); SIV-MND (mandrill) (V); and SIV-Sykes (II). Only a few representatives of the first three types are shown for sake of simplicity. Some very interesting SIV-AGM sequences from west African green monkeys, as well as some HIV-2 sequences from rural Liberia, cannot be included because they have been sequenced over only a small portion of the *pol* gene (Allan et al., 1991; Hahn, 1990). Their relevance to this inquiry will be discussed below.

The tree analysis begins with alignment of homologous nucleotide sequences. By restricting attention to just single-base changes in a relatively conserved coding sequence such as the retroviral *gag* gene, we can be reasonably confident about an unambiguous alignment. The cluster analy-

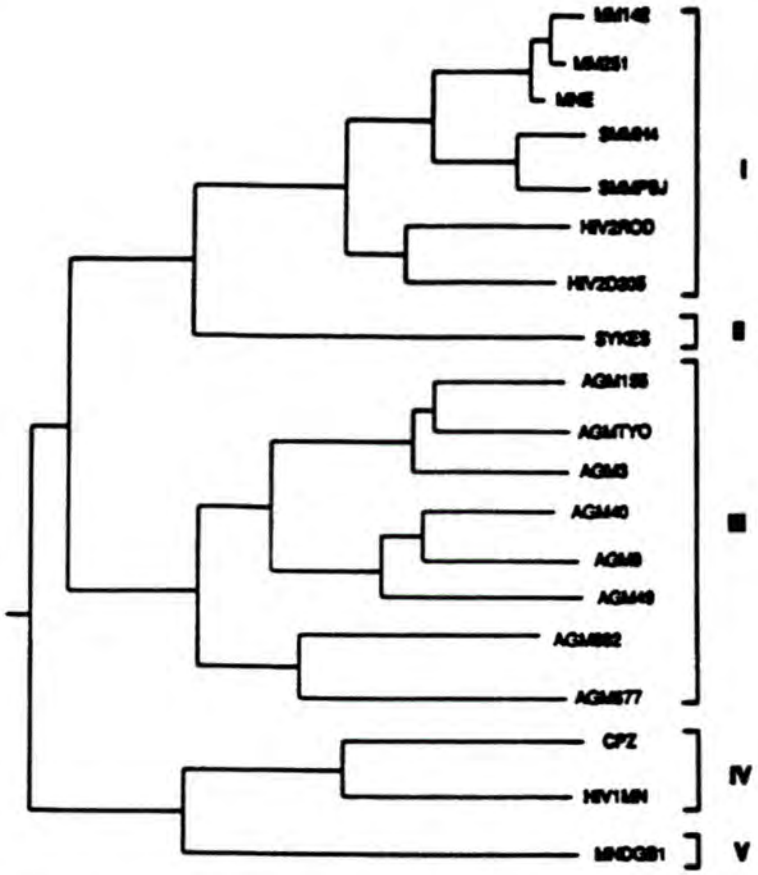


Figure 12.2 Molecular phylogenetic analysis of representative primate immunodeficiency viral gag p24 gene sequences: I, HIV-2 and related SIV's, in particular SIV-smm from the sooty mangabey; II, SIV-Sykes; III, SIV-AGMs (grivet, vervet, and tantalus species); IV, HIV-1 and the closely related SIV-cpz from a Gabonese chimpanzee; V, SIV-mnd (mandrill isolate). For the sake of clarity, only a small number of the known HIV-1 and HIV-2 sequences are shown; these are highlighted. The phylogenetic tree was generated from 648 homologous nucleotide sites (of which 416 were variable), as described in Hirsch et al. (1989) and Myers et al. (1991). Only the horizontal branch lengths should be used with the bar scale to estimate percent differences. The tree was rooted at the midpoint of the greatest patristic distance (PAUP, version 2.4.2).

sis program takes as input the alignment and infers the minimal evolutionary path for the set of sequences. The "parsimony" assumption (that the minimal evolutionary path has been followed) is appropriate for conditions of rapid mutation and recent divergence from some common ancestry. By conducting similar analyses for the several major genes of these primate immunodeficiency viruses, we can corroborate the results. That is to say, in absence of genetic recombination, a phylogenetic tree for the *pol* gene

sequences will be congruent with a tree for *gag* gene sequences. Branch lengths will vary, of course, because differing selective forces are at play, but the overall topology of the trees should be invariant (excepting recombination). The branching order can be further evaluated by a statistical procedure known as "bootstrap" analysis (Felsenstein, 1988).

As is apparent from Figure 12.2, phylogenetic tree analysis of the five known types of viruses causing human and simian AIDS implies a common ancestor for these viruses and strongly supports the hypothesis of the simian origin of AIDS. Particular attention is called to the HIV-2 and related SIV sequence group, or clade. Recently, a viral sample taken from an asymptomatic person living in rural Liberia has been found to map with the *pol* gene sequences of the mangabey (not shown in Fig. 12.2; Hahn, 1990), providing strong evidence that immunodeficiency viruses can be transmitted to humans from nonhuman primates. In light of this molecular evidence, the theory of a simian origin of AIDS has been gaining widespread acceptance (Daniel et al., 1988; Hirsch et al., 1989; Doolittle, 1989; Doolittle et al., 1990; Huet et al., 1990; Desrosiers, 1990; Gilks, 1991; Karpas, 1990; Johnson et al., 1991; Myers et al., 1992).

Although the overall picture presented in Figure 12.2 agrees in general with the cluster analyses of other investigators, some of which have been based upon a different algorithm, or even upon amino acid rather than nucleotide sequences (Smith et al., 1988; Yokoyama et al., 1988; Sharp and Li, 1988; Tsujimoto et al., 1989; Doolittle et al., 1990; Gojobori et al., 1990; Miura et al., 1990; Johnson et al., 1991), the exact branching order of the various types of primate immunodeficiency viruses whose sequences have been analyzed through cladistic analysis has encountered some ambiguities, depending upon the mode of analysis and the coding sequence under investigation. For example, the tree presented in Figure 12.2 links the SIV-AGM sequences more closely with the HIV-2's than with the HIV-1's, while only the chimpanzee viral sequence is closely linked to the type 1 HIV sequences. Furthermore, the chimpanzee viral sequence is not so closely related to the type 1 human virus as the mangabey virus is to the type 2 human viruses. Interestingly, one human viral isolate, an HIV-1 from Cameroon, is known to be more divergent from all other HIV-1's than is the SIV-cpz in Figure 12.2 (DeLeys et al., 1990; Vanden Haesevelde et al., 1991).

Doolittle's results (Doolittle, 1989; Doolittle et al., 1990), on the other hand, suggest a closer relationship between SIV-AGM sequences and HIV-1, and he has taken this result (as noted above) to support the hypothesis of an origin of HIV-1's from the green monkey reservoir. Tree analyses are typically performed on *gag* and *pol* coding sequences or amino acid sequences. Other modes of cluster analysis based upon envelope and regulatory protein sequences argue for a closer relationship between the AGMs and the HIV-2/

SIVs (Gojobori et al., 1990; Myers et al., 1992). With the recently discovered SIV-Sykes and SIV-AGM *tantalus* and *sabaeus* variants available for reevaluation of this question, not to mention the baboon and chimpanzee viral isolates now being sequenced, it should soon be possible to resolve much of this ambiguity.

A "BIG BANG" OF AIDS?

Attention should be drawn to the fact that phylogenetic tree analysis seems to suggest that there was a single point of radiation for the SIVs and thus ultimately for the HIVs that emerged from them. This may imply a special and decisive phylogenetic moment in the history of AIDS, what might be called a "big bang": it is as if several major types of the primate immunodeficiency viruses represented in Figure 12.2 emerged simultaneously and perhaps recently from a common ancestor. Most analyses to date, irrespective of the exact branching order of the primate immunodeficiency viruses, have implied this single point of radiation. This has been schematized in Figure 12.3 (Sharp and Li, 1988; Tsujimoto et al., 1989; Doolittle et al., 1990; Johnson et al., 1991).

Is the "big bang of AIDS" real or merely an artifact of the sequence analysis? Is the single point of radiation shown in Figure 12.3 actual or virtual? There is some reason to believe that the convergence of the evolutionary lines on a single point could be merely a consequence of "mutational saturation," or "noise accumulation" (Kirchhoff et al., 1990). The high rate of nucleotide substitution documented in Table 12.1 (below) implies that multiple events, creating so-called "character-conflict", and back-mutation are likely to be complicating the computer analysis, thereby casting a shadow over the distal evolutionary record of the primate lentiviruses. The unusual base composition — high adenine, low cytosine content — of HIV's and SIV's further complicates phylogenetic analyses of the "deep branches" (Sidow and Wilson, 1990). The situation has some formal parallel with the current uncertainty surrounding the theory of a cosmological big bang: radical homogeneity (mutation rates and energy radiation) and radical heterogeneity (different genomic features and galactic distributions) are both observed, and the final interpretation of the respective first moments awaits some definitive insight or discovery. Fortunately, viral relics are easier to come by than cosmic relics, and for this reason a reasonably complete understanding of the evolution of AIDS viruses may be imminent.

While it cannot be said yet whether or not all AIDS-related viruses have a single common ancestor, there is evidence that their emergence was, on an evolutionary timescale sudden and dramatic, a first decisive moment in the emergence of AIDS. For instance, descendants of African green monkeys brought to the Caribbean as early as the sixteenth century are not seropositive (Hendry et al., 1986). To our knowledge, no feral hosts for the simian

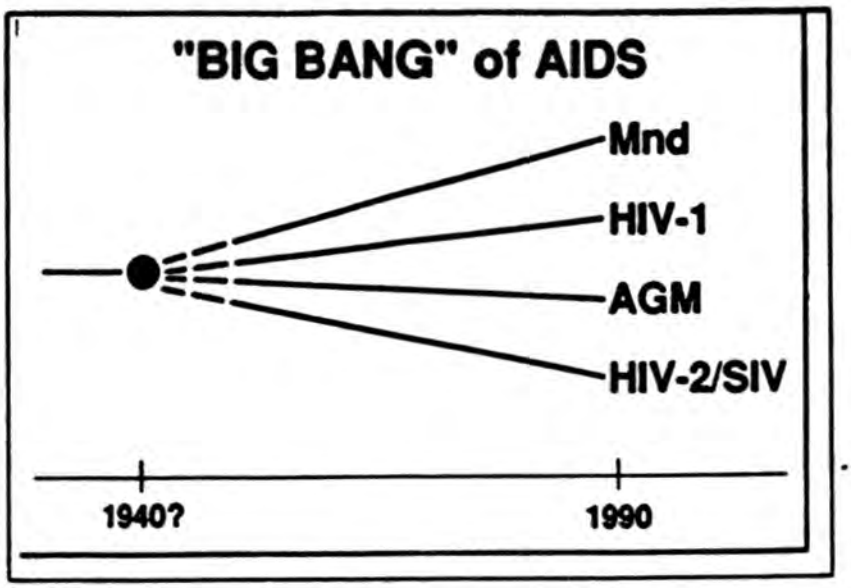


Figure 12.3 Hypothesis of a single nonhuman primate ancestral virus for four of the currently recognized five types of immunodeficiency viruses. The results shown in Figure 12.2, as well as results based upon several different analyses (see text), are suggestive of a common point of evolutionary radiation, that cannot be more recent than 50 years ago, the minimum lookback time for HIV-1 divergence from HIV-2. Alternatively, this point of radiation may be merely virtual, the consequence of mutational saturation and/or nonhomogeneous rates of nucleotide substitution (as discussed in the text).

immunodeficiency viruses have been found in South America or Asia. Only captive macaques, for example, have been found to be infected, and these undoubtedly contracted the virus from captive mangabeys. At least eight infected African nonhuman primate species are known to exist for which sequence data is still to be obtained and analyzed, among which are the baboons, colobus monkeys, guenons, and talapoins (Johnson et al., 1991).

In the third section of this paper we concluded with some certainty that the human AIDS-causing viruses have a simian origin, and in this section suggested with less certainty but significant probability that the first simian AIDS-causing viruses emerged sometime within the last few hundred years, that is quite suddenly on the scale of evolutionary time. If this is the case, what light does it shed upon the two questions that we posed at the end of the second section, namely the suddenness and virtual simultaneity of the onset of the AIDS epidemic in its major centers less than 2 decades ago, and the apparent simultaneity of the emergence of HIV-1 disease and HIV-2 disease? In order to address these questions more precisely, we will turn our attention in the next section to the time frame for the recent evolutionary development of the human AIDS-causing viruses.

DATING THE PHYLOGENETIC TREE

Tree analyses, such as Figure 12.2, do not provide explicit information about the time required for the evolutionary process being mapped by the molecular sequences. Additional hypotheses are needed. One way to estimate the time implicit in a tree is to assume a representative mutation rate, for example, a typical retroviral rate (known to be at least a millionfold higher than the rate for mammalian nuclear genes). Alternatively, the tree can be calibrated or "resolved" through the inclusion of samples acquired at different times. Both strategies have been adopted by various investigators, and a brief summary of the findings as they pertain to the divergence of HIV-1 from HIV-2 are given in Table 12.1.

The analytical results of sequence analysis performed by several groups using different strategies and different viral sequences have been in remarkable agreement. All of the estimates in Table 12.1, however, presuppose a molecular clock: reasonably uniform rates of nucleotide substitution observed over a relatively short time period—5 to 40 years, approximately—are assumed to apply in the more more distant past. There is some reason to think that this assumption may not be valid in relation to some of the feral SIV's (Allan et al., 1991; Myers et al., 1992). Moreover, if the phenomenological rate constant is a function of the number of infected individuals—"epidemic-driven variation"—then the actual divergence time could be much further back in time. In fact, by almost every other assumption, the divergence time will be greater than the estimates summarized in Table 12.1. On the other hand, even the radically lower estimate for the minimal divergence time—40 years (Fig. 12.3)—for the simian ancestors of HIV-1

Table 12.1 HIV-1/HIV-2 Divergence Time Estimates

Base Substitution Rates (per site per year)	Basis	HIV Divergence Time (minimum years)	Reference
0.0005	Murine retroviral gag genes	280	Yokoyama et al., 1988
0.0116	HIV-1/2 env genes	40	Smith et al., 1988
0.001	HIV pol genes	150	Sharp and Li, 1988
0.0008	Ovine lentiviral pol genes	203	Querat et al., 1990

and HIV-2 does not help us understand why these viruses first appeared in humans only within the last 2 decades (as discussed in the section Epidemiological Snapshots, above). Furthermore, the phylogenetic record does not provide a satisfactory explanation for the enigma of two virtually simultaneous cross-species transmissions.

DUAL CROSS-SPECIES TRANSMISSION

Can any more be said at this time about the second decisive moment in the AIDS pandemic, the twofold introduction of virus into human hosts? Assuming (with greater confidence now) a simian origin for the AIDS viruses, and acknowledging the improbability that the two independent and virtually simultaneous cross-species transmissions took place naturally (i.e., without human intervention), we will consider several contexts in which human accients might have occurred and that could explain the coincidence that we have been calling, for lack of better description, the twofold epidemic. First, there was a sharp increase in the exportation of monkeys from Africa in the 1960s, largely for purposes of medical research. This entailed extensive handling of captured animals (Conner and Kingman, 1988). Second, there was also in the 1960s a valiant push to vaccinate Third World populations. In particular, many countries began using live polio vaccines in the spring of 1960, creating some small but finite opportunity for introduction of passenger viruses through contaminated vaccine lots, as in fact happened with SV-40. A great potential for amplification of infections, natural or accidental, exists in the use of contaminated needles.

Short of discovering SIV-infected lots of vaccines prepared in the 1960s, it seems unlikely that a clear verdict regarding these possible modes of transmission will ever come forward. Blomberg and coworkers (1990) have recently documented cross-reactivities of normal human sera with capsid proteins of diverse simian retroviruses, suggesting a way to evaluate some of these potential modes of transmission. Such antibodies probably have arisen as a result of accidental exposure to simian retroviral antigens, although molecular mimicry cannot be ruled out. To the extent that the putative twofold cross-species transmission resulted from increased handling of animals and use of unsterile needles, there exists the disturbing prospect of future transmissions out of the large, qualitatively diverse pools of simian immunodeficiency viruses that are now being discovered.

There is, to keep the record straight, no evidence supporting any of the human accident hypotheses beyond the circumstantial evidence outlined above. And the possibility of uncovering long-established but highly isolated endemic sources in humans—the so-called "lost tribe" hypothesis—for which Hahn (1990) has provided a bit of evidence in relation to the HIV-2 epidemic in west Africa, cannot be dismissed at this time.

MOLECULAR EPIDEMIOLOGY

Sampling and sequencing of HIV-1 up to the present has provided a surprisingly coherent geographical picture of the viruses responsible for the major part of the pandemic: envelope gene sequences from Zairean samples have been clustering together and sequences from U.S. isolates have been clustering together. This neat segregation of viral forms, which is testimony to the "newness" of AIDS, is doomed to be transitory. By 1992, at least 5 distinct subtypes of HIV-1 could be identified, based upon *gag* and *env* coding sequences, in major centers of the pandemic. At least 2 of these 5 forms of HIV-1 are now known to be cocirculating in Thailand, and another two forms in Brazil; thus, it is only a matter of years, at most a decade, before a redistribution of HIV-1 variants is achieved. In the face of the extreme diversity of HIV-1 and HIV-2, and because of the limited sampling thus far (until recently all African HIV-1 samples had been taken from Kinshasa), several international programs of molecular epidemiology have been mounted; now we have sequenced samples from Gabon, Cameroon, Uganda, Kenya, and South Africa, to name some of the sites under study. The results of these programs will be especially important for the vaccine trials that are anticipated.

The random sampling and sequencing to date—molecular snapshots so to speak from HIV-1 samples representing some of the earliest infections—presents a sobering picture of what has to be expected in the future from these emerging viruses. These sequence studies have revealed that viruses with envelope gene sequences differing by 0% to 6% are typical of clones from a single patient (intrapatient variation), while viruses with sequences differing by 6% to 12% have been characteristic of samples from different patients (interpatient variation) in the United States up to 1985. In the 1990-1991 investigation of a Florida dentist and the patients to whom he apparently transmitted virus in the course of invasive oral procedures (Morbidity and Mortality Weekly Report of 18 January, 1991, published by the U.S. Centers for Disease Control), viral samples at large in the Florida population were found to differ by up to 19%.

This more recent sampling of descendants of the earlier forms in the United States (there is little evidence of significant migration of new forms into the United States between 1985 and 1990) implies that the average interpatient distance appears to be increasing at approximately 1% per year. This rough calculation is in fairly good agreement with the envelope nucleotide substitution rates reported in several recent studies involving other cohorts as well as single individuals (reviewed in Myers and Pavlakis, 1992). Since it is inevitable that both viral migration and recombination will be making significant contributions to the total diversity over a longer period of the pandemic, the value of 1% per year for single base changes should be regarded as a minimum rate.

There is no clear evidence to date that individuals can be dually or multiply infected with the same type of immunodeficiency virus; if super-

infection can naturally occur, it should soon become apparent with the Thai and Brazilian samples that are now under investigation. HIV-1 envelope gene sequences that differ by 40% are known to exist for highly divergent viruses found in the United States and Africa. The extraordinary variation of AIDS viruses, observed especially in human and captive animal populations, may be a paradigmatic example of non-steady state evolution, or "bottom-heavy cladism" resulting from a recent ecological breakthrough (as stated by Gould et al., 1987, "Clades diversify rapidly in ephemeral times of unusual opportunity").

Tracking human immunodeficiency viruses is not a futile task. Precisely because these retroviruses vary so rapidly, we discover all the more quickly the conserved and invariant properties of their genes and proteins (Myers et al., 1991). The design of antiviral agents will be enormously facilitated by the fact that these viruses, unlike more conservative microbes, are unable to keep the essential elements of their genomes a secret.

SUMMARY AND CONCLUSION

The preceding discussion has suggested the following results:

1. Phylogenetic tree analysis points strongly to a simian origin of the human immunodeficiency viruses. This conclusion is corroborated by cluster analyses done by many other investigators.

2. The five (or more) major types of viruses that infect simians and cause AIDS in humans may have emerged from a common ancestor at about the same time. This arguably first phylogenetic moment of the AIDS epidemic, which may be likened to the "big bang", cannot explain the sudden and concurrent appearance of human AIDS in the known centers of the epidemic, beyond accounting for a very large and diverse reservoir of nonhuman primate immunodeficiency viruses present in Africa at the onset of the pandemic.

3. The divergence of the type 1 and type 2 human AIDS viruses may have taken place as recently as 40 years ago, though most estimates would place it at 150 to 250 years ago. While this is a relatively short time on an evolutionary scale, even the minimum 40-year figure is too long to explain the sudden and concurrent emergence of the epidemic in its major centers only as recently as the late 1970s.

4. In brief, neither the hypothetical "big bang" nor the range of possible dates for divergence of HIV-1 from HIV-2 accounts for the "improbable" simultaneous cross-species transmission of both HIV-1 and HIV-2 that apparently occurred.

5. These perplexing facts might be explained by two possible avenues of cross-species transmission that occurred in the 1960s: the widespread use in Africa of needles in association with vaccination programs (and

possibly of live virus vaccines cultured on monkey tissue), and the sharp increase in exportation of monkeys from Africa, largely for medical research, which entailed extensive handling of captured monkeys. There is no direct evidence for either of these possibilities, however the transmission of virus in 1988 from an infected dentist to five of his patients painfully reminds us that human accidents can be primary avenues for viral emergence.

6. Sampling and sequencing projects now underway will provide additional information about variations in the nucleotide and protein sequences of the AIDS-causing viruses and their rate of change. This will throw light not only on the evolutionary history of these viruses, but on their likely future.

ACKNOWLEDGMENTS

Vanessa Hirsch and Phillip Johnson, of Georgetown University and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health in the United States, kindly provided prior to publication the sequences for the Sykes and *tantalus* viral isolates analyzed in Figure 12.2. The HIV Sequence Database and Analysis Project is funded by the AIDS Division of the National Institute of Allergy and Infectious Diseases through an interagency agreement with the U.S. Department of Energy (LAUR 89-3408).

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.

Conner, S. and S. Kingman (1988). *The Search for the Virus*. London: Penguin.

Daniel, M. D., Y. Li, Y.M. Naidu, P.J. Durda, D. K. Schmidt, C.D. Troup, D.P. Silva, J.J. MacKey, H.W. Kestler III, P.K. Sehgal, N.W. King, Y. Ohta, M. Hayami, and D.C. Desrosiers (1988). Simian immunodeficiency virus from African green monkeys. J. Virol. 62:4123-4128.

DeCock, K. M., J.C. Niland, H.P. Lu, V. Edwards, C. Shriver, J.W. Mosley, et al. (1987). Human immunodeficiency virus infection in patients with hepatitis B virus and hepatitis delta virus infections in Los Angeles, 1977-1985. Abstract MP.84, Third International Conference on AIDS, Washington, D.C., 1-5 June 1987.

De Leys, R., B. Vanderborght, M. Vanden Haesevelde, L. Heyndrickx, A. van Geel, C. Wauters, R. Bernaerts, E. Saman, P. Nijs, B. Willems, H. Taelman, G. van der Groen, P. Piot, T. Tersmette, J.G. Huisman, and H.

simplex viruses generate lethal recombinants in vivo. Science 234:746-748.

Johnson, P., G. Myers, and V.M. Hirsch (1991). Genetic diversity and phylogeny of nonhuman primate lentiviruses. In *AIDS Research Reviews*, vol. 1 (W.C. Koff, F. Wong-Staal and R.C. Kennedy, eds.), pp. 47-62. New York: Marcel Dekker.

Kanki, P.J., J. Alroy, and M. Essex (1985). Isolation of T-lymphotropic retrovirus related to HTLV-III/LAV from wild-caught African green monkeys. Science 230:951-954.

Karpas, A. (1990). Origin and spread of AIDS. Nature 348:578.

Kirchhoff, F., K. Dieter Jentsch, B. Bachmann, A. Stuke, C. Laloux, W. LHke, C. Stahl-Hennig, J. Schneider, K. Nieselt, M. Eigen, and G. Hunsmann (1990). A novel proviral clone of HIV-2: biological and phylogenetic relationship to other primate immunodeficiency viruses. Virology 177:305-311.

Marx, P.A., Y. Li, N.W. Lerche, S. Sutjipto, A. Gettie, J.A. Yee, B.H. Brotman, A.M. Prince, A. Hanson, R.G. Webster, and R.C. Desrosiers (1991). Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a west African pet sooty mangabey. J. Virol. 65:4480-4485.

Miura, T., J.-i. Sakuragi, M. Kawamura, M. Fukasawa, E.N. Moriyama, T. Gojobori, K.-i. Ishikawa, J.A.A. Mingle, V.B.A. Nettey, H. Akari, M. Enami, H. Tsujimoto, and M. Hayami (1990). Establishment of a phylogenetic survey system for AIDS-related lentiviruses and demonstration of a new HIV-2 subgroup. AIDS 4:1257-1261.

Myers, G., and G.N. Pavlakis (1992). Evolutionary potential of complex retroviruses. In *The Retroviridae* (J.A. Levy, ed.), vol. 1. New York: Plenum.

Myers, G., B. Korber, J.A. Berzofsky, R.F. Smith, and G.N. Pavlakis (1991). *Human Retroviruses and AIDS 1991*. Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, N.M. 87545.

Myers, G., K. MacInnes, and B. Korber (1992). The emergence of simian/human immunodeficiency viruses. AIDS Res. Hum. Retroviruses 8:.

Nahmias, A.J., J. Weiss, X. Yao, F. Lee, R. Kodsi, M. Schanfield, T. Matthews, D. Bolognesi, D. Durack, A. Motulsky, P. Kanki, and M. Essex (1986). Evidence for human infection with an HTLV-III/LAV-like virus in central Africa. Lancet i:1279-1280.

Palca, J. (1991) The sobering geography of AIDS [News Article]. Science 252:372.

Pape, J.W., M.E. Stanback, M. Pamphile, R. Verdier, M.-M. Deschamps, W.D. Johnson Jr., et al. (1987). Pattern of HIV infection in Haiti: 1977-1986. Abstract M.8.6, Third International Conference on AIDS, Washington, D.C., 1-5 June 1987.

Piot, P., F.A. Plummer, M. Rey, E.N. Ngugi, C. Rouzioux, J.O. Ndinya-Achola, L.J. D'Costa, G. Vercauteren, et al. (1986). Retrospective seroepidemiology of AIDS virus infection in Nairobi populations. Abstract S10a, Second International Conference on AIDS, Paris, 23-25 June 1986.

Piot, P., F.A. Plummer, F.S. Mhalu, J.-L. Lamboray, J. Chin, and J.M. Mann (1988). AIDS: an international perspective. Science 239:573-579.

Querat, G., G. Audoly, P. Sonigo, and R. Vigne (1990). Nucleotide sequence analysis of SA-OMVV, a visna-related ovine lentivirus: phylogenetic history of lentiviruses. Virology 175:434-447.

Rayfield, M., K. DeCock, W. Heyward, L. Goldstein, J. Krebs, S. Kwok, S.

Lee, J. McCormick, J.M. Moreau, K. Odehouri, G. Schochetman, J. Salasky, and C.-Y. Ou (1988). Mixed human immunodefiency virus (HIV) infection in an individual: demonstration of both HIV type 1 and type 2 proviral sequences by using polymerase chain reaction. J. Infect. Dis. 158:1170-1176.

Rouzioux, Ch., G. Jaeger, F. Brun-Vezinet, M.C. Dazza, M.A. Rey, L. Montagnier, G. Charmot, et al. (1986). Absence of antibody to LAV/HTLV-III and STLV-III (mac) in pygmies. Abstract 378, Second International Conference on AIDS, Paris, 23-25 June 1986.

Sharp, P. M., and W.-H. Li (1988). Understanding the origins of AIDS viruses. Nature 336:315.

Sher, R., E. Antunes, B. Reid, and H. Falcke (1987). HIV antibody prevalence in black miners between 1970-1974. Abstract MP.42, Third International Conference on AIDS, Washington, D.C., 1-5 June 1987.

Sidow, A., and A.C. Wilson (1990). Compositional statistics: an improvement of evolutionary parsimony and its application to deep branches in the tree of life. J. Mol. Evol. 31:51-68.

Smith, T. F., A. Srinivasan, G. Schochetman, M. Marcus, and G. Myers (1988). The phylogenetic history of immunodeficiency viruses. Nature 333:573-575.

Srinivasan, A., D. York, D. Butler Jr., R. Jannoun-Nasr, J. Getchell, J. McCormick, C.-Y. Ou, G. Myers, T. Smith, E. Chen, G. Flaggs, P. Berman, G. Schochetman, and S. Kalyanaraman (1989). Molecular characterization of HIV-1 isolated from a serum collected in 1976: nucleotide sequence comparison to recent isolates and generation of hybrid HIV. AIDS Res. Hum. Retroviruses 5:121-129.

Stevens, C.E., P.E. Taylor, E.A. Zang, J.M. Morrison, E.J. Harley, S.R. de Cordoba, C. Bacino, R.C.Y. Ting, A.J. Bodner, M.G. Sarngadharan, R.C. Gallo, and P. Rubinstein (1986). Human T-cell lymphotropic virus type III infection in a cohort of homosexual men in New York City. JAMA 255:2167-2172.

Titti, F., A. Aceti, P. Verani, G.B. Rossi, K. Bile, and A. Sebastiani (1986). Seroepidemiologic screening for antibodies to HTLV-III/LAV in Somalia. Abstract 380, Second International Conference on AIDS, Paris, 23-25 June 1986.

Tsujimoto, H., A. Hasegawa, N. Maki, M. Fukasawa, T. Miura, S. Speidel, R.W. Cooper, E.N. Moriyama, T. Gojobori, and M. Hayami (1989). Sequence of a novel simian immunodeficiency virus from a wild-caught African mandrill. Nature 341:539-541.

Vanden Haesevelde, M., J.L. Decourt, L. Heyndricks, B. Vanderborght, E. Saman, R. DeLeys, G. van der Groen, and H. Van Heuverswyn (1991). Molecular cloning and complete sequence analysis of a highly divergent African HIV isolate. Abstract M.A.1157, Seventh International Conference on AIDS, Florence, 16-21 June 1991.

Yokoyama, S., L. Chung, and T. Gojobori (1988). Molecular evolution of the human immunodeficiency and related viruses. Mol. Biol. Evol. 5:237-251.

Van Heuverswyn (1990). Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. J. Virol. 64:1207-1216.

Desmyter, J., P. Coubau, S. Chamaret, and L. Montagnier (1986). Anti-LAV/HTLV-III in Kinshasa mothers in 1970 and 1980. Abstract S17g, Second International Conference on AIDS, Paris, 23-25 June 1986.

Desrosiers, R. C. (1990). A finger on the missing link. Nature 345:288-289.

Doolittle, R. F. (1989). Immunodeficiency viruses: The simian-human connection. Nature 339:338-339.

Doolittle, R. F., D.-F. Feng, M.A. McClure, and M.S. Johnson (1990). Retrovirus phylogeny and evolution. In *Retroviruses. Strategies of Replication* (R. Swanstrom and P.K. Vogt, eds.), pp. 1-18 [Current Topics in Microbiology and Immunology, 157]. Berlin: Springer-Verlag.

Duesberg, P. H. (1989). Human immunodeficiency virus and acquired immunodeficiency syndrome: Correlation but not causation. Proc. Nat. Acad. Sci. U.S.A. 86:755-764.

Felsenstein, J. (1988) Phylogenies from molecular sequences: Inference and reliability. Ann. Rev. Genet. 22:521-565.

Fultz, P. N., W.M. Switzer, C.A. Schable, R.C. Desrosiers, D. Silva, and J.B. McCormick (1987). Serologic evidence for infection by HIV-2 in Guinea Bissau in 1980. Abstract MP.72, Third International Conference on AIDS, Washington, D.C., 1-5 June 1987.

Getchell, J.P., D.R. Hicks, A. Srinivasan, J.L. Heath, D. A. York, M. Malonga, D.N. Forthal, J.M. Mann, and J.B. McCormick (1987). Human immunodeficiency virus isolated from a serum sample collected in 1976 in central Africa. J. Infect. Dis. 156: 833-837.

Gilks, C. (1991). AIDS, monkeys and malaria. Nature 354:262.

Gojobori, T., E.N. Moriyama, Y. Ina, K. Ikeo, T. Miura, H. Tsujimoto, M. Hayami, and S. Yokoyama (1990). Evolutionary origin of human and simian immunodeficiency viruses. Proc. Natl. Acad. Sci. USA 87:4108-4111.

Gould, S. J., N.L. Gilinsky, and R.Z. German (1987). Asymmetry of lineages and the direction of evolutionary time. Science 236:1437-1441.

Hahn, B. H. (1990). Biologically unique SIV-like HIV-2 variants in healthy west African individuals. In *Retroviruses of Human A.I.D.S. and Related Animal Diseases* [Proceedings of the Fifth Cent Gardes Colloquium, Paris, Oct. 29-31, 1990] (M. Girard and L. Valette, eds.), pp. 31-38. Paris: Pasteur-Merieux.

Hendry, R.M., M.A. Wells, M.A. Phelan, A.L. Schneider, J.S. Epstein, and G.V. Quinnan (1986). Antibodies to simian immunodeficiency virus in African green monkeys in Africa in 1957-62. Lancet ii:455.

Hirsch, V. M., R.A. Olmsted, M. Murphey-Corb, R.H. Purcell, and P.R. Johnson (1989). An African primate lentivirus (SIV_{sm}) closely related to HIV-2. Nature 339:389-392.

Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson (1990). Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature 345:356-359.

Jaffe, H.W., W.W. Darrow, D.F. Echenberg, P.M. O'Malley, J.P. Getchell, V.S. Kalyanaraman, R.H. Byers, D.P. Drennan, E.H. Braff, J.W. Curran, and D.P. Francis (1985). The acquired immunodeficiency syndrome in a cohort of homosexual men. Ann. Int. Med. 103:210-214.

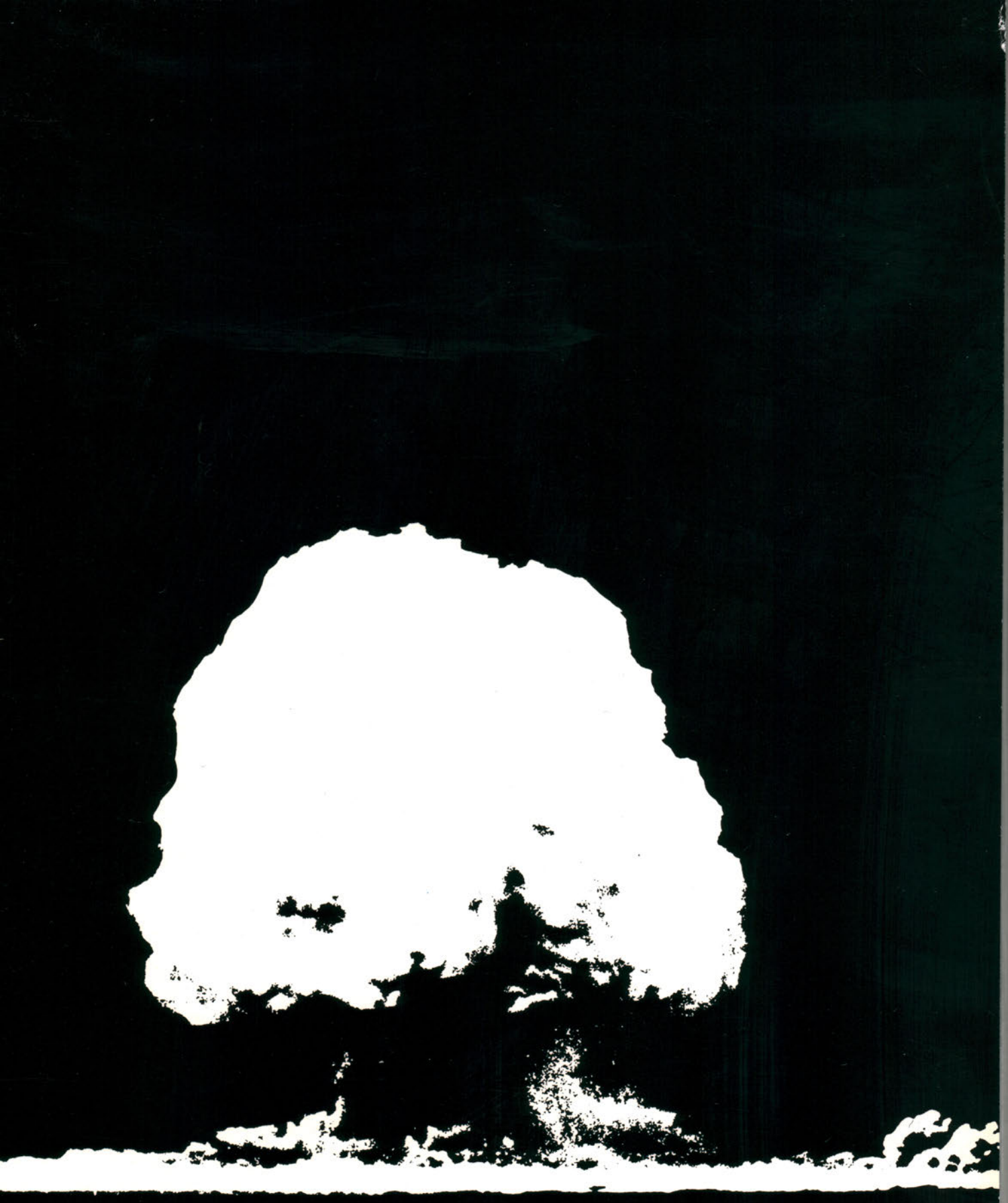
Javier, R.T., F. Sedarati, and J.G. Stevens (1986). Two avirulent herpes

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.



Los Alamos 1943-1945; The Beginning of an Era