When HIV spread afar

Edward C. Holmes*

Department of Biology, Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, PA 16802; and Fogarty International Center, National Institutes of Health, Bethesda, MD 20892

Since the first cases of AIDS were described in the United States in 1981, the origin of this devastating disease has intrigued scientists and the general public alike. This fascination is reflected in the numerous theories put forth for the emergence of HIV, the most infamous of which involves the alleged use of HIV-contaminated oral polio vaccine in Africa during the late 1950s (1). Thankfully, a steady stream of virological data and phylogenetic analyses now means that the oral polio vaccine theory has rightly been assigned to the back shelves of science fiction (2). The article by Gilbert *et al.* in this issue of PNAS (3) similarly uses an elegant combination of virology and phylogeny to shed light on another key moment in the history of HIV: its spread from an origin in Africa to the Americas.

Haiti: Sink or Source?

From the earliest days of AIDS reporting, it was clear that the Caribbean nation of Haiti was particularly significant in this epidemic. Indeed, HIV/AIDS was initially found to be relatively frequent in persons of Haitian origin, and some Haitian isolates of HIV-1 fell on relatively deep branches in phylogenetic trees, suggesting that the virus took an early foothold in that country. Until now, however, the connection among Africa, Haiti, and industrialized nations like the United States has largely remained the stuff of speculation. By deploying an impressive armory of phylogenetic techniques, Gilbert *et al.* (3) provide the first solid evidence for the role of Haiti in the emergence and evolution of HIV.

Like many RNA viruses and retroviruses, HIV-1 is genetically very diverse, falling into a series of phylogenetically defined clades, or subtypes, that have differing geographic distributions, as well as an ever-expanding set of intersubtype recombinants. The focus of this particular study is HIV-1 subtype B, the form of virus that was first described in U.S. populations in the early 1980s and that still dominates infections in most industrialized nations to the present day. The global spread of subtype B is considered a major event in the history of HIV/AIDS because it marks the point when the virus first entered the large, wealthy, and highly mobile populations of the Western world.

Two factors contribute to the power of the Gilbert *et al.* (3) study: (*i*) the use of sophisticated methods of sequence analysis that are able to account for some of the idiosyncrasies of HIV evolution and (*ii*) the retrieval of gene sequence data from "archival" HIV samples, notably those from patients of Haitian origin who carried the virus in the early 1980s. Although far older samples are available from a number of other viruses (for example, those for human influenza A virus date back to 1918; ref. 4), these HIV viruses are certainly old with respect to the spread of HIV outside of Africa and so provide a unique window into the timescale of viral evolution. With this happy marriage of new sequence data and state-of-the-art bioinformatics, Gilbert *et al.* first show that those subtype B viruses in Haiti have their origins in Africa. They then provide compelling evidence that Haiti has unwittingly acted as the conduit for the spread of HIV to the United States and a wide range of other localities rather than being simply a regional sink.

Of more interest are the attempts of Gilbert *et al.* (3) to put these evolutionary events into an historical time frame. To achieve this result, the authors estimated the time to the most recent common ancestor (TMRCA) of HIV-1 subtype B using a "relaxed" molecular clock (5), which allows the rate of evolutionary change to vary in a lineage-specific manner, a major factor in the evolution of HIV. They estimated that the date for the spread of HIV-1 to Haiti from its ancestry in Africa is between 1962 and 1970 (with a mean of 1966). Importantly, this timescale corresponds well with a period when many Haitians returned to their home country from the Congo, after the latter's independence from Belgium and subsequent political crises. Because the Congo region has been shown to play a pivotal role in the genesis of HIV (6), the correspondence between the travel data and the inferred epidemiological timescale of Gilbert *et al.* provides strong circumstantial evidence that the timescale is broadly correct. This study therefore highlights the role played by socioeconomic factors such as human migration in the history of infectious disease. In addition, the migration of HIV from Haiti to the United States and beyond is dated to the period 1966–1972, perhaps 30–40 years after the virus first established itself in the human population in Africa.

A Slow Fuse for AIDS in the Americas?

Perhaps the most fascinating insight from this exercise in viral archeology is that HIV was spreading in the United States for at least 9 years before its first clinical description. This finding is bound to spark a lively debate because it may seem untenable that such a long period of "cryptic" transmission could be possible in a nation with such an advanced health care system. Some may argue that the now-characteristic symptoms of severe immunodeficiency would have been spotted sooner, even given the time lag between initial infection and the onset of AIDS. Although it is theoretically possible that the high virulence of HIV infection—manifest as AIDS—did not evolve until later in the U.S. epidemic, such that HIV virulence has increased through time, a far more likely explanation for a long history of HIV in the United States before the discovery of "patient zero" is simply that it went undetected or was misdiagnosed. Indeed, this cryptic period is far shorter than an equivalent period in Africa, where the disease may have remained unrecognized for more than half a century. Furthermore, for the "increasing virulence" hypothesis to be true, all of the subtypes of HIV-1 that have independent origins in central-west Africa would have to have evolved the symptoms of AIDS independently, which seems untenable. A slow fuse for the explosion of AIDS in the United States in the early 1980s is also compatible with serological studies that suggest that thousands of individuals may have already been HIV-infected in this country by the late 1970s (7).

Another possibility is that the relaxed molecular clock used by Gilbert *et al.* (3), although a major advance, does not fully capture the nature of HIV evolution. The most likely cause of any clock error is that HIV exhibits rather differ-

Author contributions: E.C.H. wrote the paper.

The author declares no conflict of interest.

See companion article on page 18566.

*E-mail: ech15@psu.edu.

ent evolutionary rates at the intrahost and interhost levels. In particular, there is mounting evidence for an inverse relationship between rates of viral transmission and rates of evolutionary change, with the highest rates observed within individual hosts (8). Consequently, the very rapid spread of HIV through standing networks of gay men and injecting drug users in industrialized nations during the early 1980s may have been characterized by unusually low rates of evolutionary change, which in turn will introduce error into estimates of the TRMCA. This important dynamical relationship has two possible causes: (*i*) that intrahost evolution, in contrast to that occurring among hosts, is dominated by the positive selection of amino acid changes that facilitate immune escape and that elevate rates of evolutionary change over that expected under neutral genetic drift (9), and/or (*ii*) that most of the mutations that occur within hosts are purged at transmission to new hosts because of strong purifying selection in this new environment. For example, a proportion of the HIV genome appears to be "reset" at interhost transmission because of mismatches between mutations that confer escape from host cytotoxic T lymphocyte responses and the HLA type determining the specificity of that response (10). In short, mutations that are advantageous in one individual may be detrimental in another. This continual rewinding may slow the molecular clock in fast epidemics.

The global spread of subtype B is considered a major event in the history of HIV/AIDS.

Although it is possible that the timescale for HIV evolution proposed by Gilbert *et al.* (3) has, to some extent, been adversely affected by changing rates of epidemic spread, the correspondence between the documented movement of individuals from Africa to Haiti and the dates estimated in this paper make it likely that any rate variation is adequately encompassed within the distribution of evolutionary rates estimated under a relaxed molecular clock. While conclusive proof for the cryptic transmission of HIV will obviously require the sampling, sequencing, and phylogenetic analysis of HIV samples from the United States obtained during the 1970s, this paper undoubtedly sets the benchmark for future studies in viral phylogeography.

1. Worobey M, Santiago ML, Keele BF, Ndjango J-BN, Joy JB, Labama BL, Dhed'a BD, Rambaut A, Sharp PM, Shaw GM, Hahn BH (2004) *Nature* 428:820.
2. Hooper E (1999) *The River: A Journey Back to the Source of HIV and AIDS* (Little, Brown, Boston).
3. Gilbert MTP, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M (2007) *Proc Natl Acad Sci USA* 104:18566–18570.
4. Taubenberger JK, Reid AH, Frafft AE, Bijwaard KE, Fanning TG (1997) *Science* 275:1793–1796.
5. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) *PLoS Biol* 4:e88.
6. Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC (2001) *Nature* 410:1047–1048.
7. Jaffe HW, Darrow WW, Echenberg DF, O'Malley PM, Getchell JP, Kalyanaraman VS, Byers RH, Drennan DP, Braff EH, Curran JW, *et al.* (1985) *Ann Intern Med* 103:210–214.
8. Berry IM, Ribeiro R, Kothari M, Athreya G, Daniels M, Lee HY, Bruno W, Leitner T (2007) *J Virol* 81:10625–10635.
9. Nielsen R, Yang Z (1998) *Genetics* 148:929–936.
10. Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, Kelleher AD, Allen TM (2007) *J Virol* 81:193–201.