J Mol Evol (2003) 56:643–644
DOI: 10.1007/s00239-002-2438-6

JOURNAL OF MOLECULAR EVOLUTION

Letters to the Editor

Recent Evolutionary History of Human Immunodeficiency Virus Type 1 Subtype B—Response

Una R. Smith, Carla Kuiken, Bette T. Korber

Theoretical Biology and Biophysics, Los Alamos National Laboratory, MS K-710, Los Alamos, NM 87545, USA

Received: 9 September 2002 / Accepted: 18 December 2002

Abstract. The year of origin estimated by Lukashov and Goudsmit for HIV-1 subtype B is 1976 (95% CI, 1974–1977); this is significantly different from our prior estimate, 1967 (95% CI, 1960–1971). We review published evidence, which suggests that their estimate is too late.

Key words: HIV-1 — AIDS — Epidemic — Molecular clock — Most recent common ancestor

Lukashov and Goudsmit (2002) misrepresent our work (Korber et al. 2000) on the origin of the HIV-1M subtype B epidemic by mentioning only our point estimate of the year of origin, omitting the confidence interval (CI). Our best estimate, based on full-length envelope sequences, suggests that the most recent common ancestor (MRCA) of subtype B in the United States (US) originated in 1967 (95% CI, 1960–1971). It is misleading to omit the confidence interval, particularly because the distribution of the error around the point estimate is not normal.

Aside from this, the arguments Lukashov and Goudsmit present to suggest that the subtype B epidemic in the US had an MRCA in 1976 (95% CI, 1974–1977) rest on several untenable assumptions and inferences.

First, equating the time of the first confirmed AIDS case minus a few years to the time of introduction of the virus, even in the US, is unrealistic. By its nature, AIDS is difficult to confirm retrospectively, so the odds are very remote that all AIDS cases have been found. The initial estimates of the origin of HIV in the US, which Lukashov and Goudsmit use as their target date, were made at a time when the time of progression to AIDS was unknown but often assumed to be shorter than it is now known to be. For example, less than 5% of people infected with HIV progress to AIDS within 2–3 years (Mellors et al. 1996; Phair et al. 1992); these unfortunate people are known as rapid progressors. Retrospectively, 12 AIDS cases were identified in 1978–1979 (Selik et al. 1984). If those 12 people were infected in 1976 at the earliest, then they necessarily were rapid progressors and, as such, would represent 5% of at least several hundred people infected in 1976 alone. This is very difficult to reconcile with the estimate of an MRCA that same year. On the other hand, if the AIDS cases in 1978–1979 represent a typical range of progression rates, including some with long asymptomatic periods, then the MRCA of subtype B must be older than 1976. Huminer and Pitlik (1988) list other possible (but unconfirmed) AIDS cases worldwide that go back to 1953. In both Haiti and the US, scattered cases of HIV-1 infection or AIDS were confirmed in the late 1970s, with a handful of possible and probable cases between 1972 and 1976 (Hooper 1999, pp. 77–82, 440–443). By the first quarter of 1983 there

Correspondence to: Una Smith; *email*: una@lanl.gov

were already 1299 confirmed AIDS cases reported in the US, spread over 35 states (Selik et al. 1984).

Second, the estimation method used by Lukashov and Goudsmit is unsuitable for HIV. In linear regression analysis, the error of the observations around the linear fit is implicitly assigned to the variable on the *Y*-axis, in this case the divergence. But HIV biology tells us that there is significant error on the time axis as well. (a) Sample dates are usually recorded only to the nearest calendar year. (b) The sampled virus can be a recently generated variant or can represent genetic material that has been sequestered for years in a latently infected cell. This means that the sample date may differ by as much as 10 years from the evolutionary age due to latency in a patient. On a time scale of 20–30 years, that difference is enormous. Our analyses (Korber et al. 2000) showed that ignoring this asymmetric error on the time axis leads to an underestimation of the age of the epidemic.

Third, Lukashov and Goudsmit completely ignore Haiti in their estimate of the origin of the global subtype B epidemic. A scenario that is widely considered plausible is that subtype B (or a B/D ancestor) was introduced into the US via Haiti through Haitian–African contacts in the 1960s. While the first AIDS cases were found in Haiti and the US at around the same time (1976–1978), in 1978 the prevalence of HIV was found to be much higher in Haiti than in the US (Johnson and Pape 1989). Excluding Haiti thus excludes what are potentially the earliest instances (early samples and/or samples of early lineages) of subtype B, and this could seriously underestimate the age of the B-clade epidemic. Also, to estimate the global MRCA of subtype B, Lukashov and Goudsmit used sequences from just four countries, all of which probably received subtype B from the US. If so, their MRCA is not global but rather is the same as (or a descendant of) the MRCA of the US epidemic; we note that their estimates for the global and US epidemics are the same. We surmise that the true global subtype B epidemic, which must include Haiti, is older.

The simplistic approach to regression analysis, the exclusion of samples from Haiti, and the unlikely assumption that the first confirmed AIDS cases would be found within a few years from the MRCA despite HIV's 10-year average latency, all lead to an underestimation of the age of the subtype B epidemic. Lukashov and Goudsmit cite the fact that their estimate based on only silent changes in all cases is closer to their target date as an indication that this estimate is better than those including nonsilent changes, which generally result in earlier dates. Since, in our opinion, their target date is too late, it can be argued that including only silent changes makes their estimate worse, not better.

References

Hooper E (1999) The river. Little, Brown, Boston, MA

Huminer D, Pitlik SD (1988) AIDS in the pre-AIDS era. Can Med Assoc J 138:403

Johnson WD Jr, Pape, JW (1989) AIDS in Haiti. Immunol Ser 44:65–78

Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S, Bhattacharya T (2000) Timing the ancestor of the HIV-1 pandemic strains. Science 288:1789–1796

Lukashov VV, Goudsmit J (2002) Recent evolutionary history of human immunodeficiency virus type 1 subtype B: Reconstruction of epidemic onset based on sequence distances to the common ancestor. J Mol Evol 54:680–691

Mellors JW, Rinaldo CR Jr, Gupta P, White RM, Todd JA, Kingsley LA (1996) Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science 272:1167–1170

Phair J, Jacobson L, Detels R, Rinaldo C, Saah A, Schrager L, Munoz A (1992) Acquired immune deficiency syndrome occurring within 5 years of infection with human immunodeficiency virus type-1: The Multicenter AIDS Cohort Study. J Acquir Immune Defic Syndr 5:490–496

Selik RM, Haverkos HW, Curran JW (1984) Acquired immune deficiency syndrome (AIDS) trends in the United States, 1978–1982. Am J Med 76:493–500