QAD 200022

Correspondence

AIDS 2005, **19**:000–000

Time and rate of evolution are the key to establish transmission cases

Recent controversy has been expressed in this journal on the adequacy of single gene phylogenies to establish confidently true relationships in transmission cases [1,2]. The initial analysis by Stürmer *et al.* [3] of the *pol* and *env* regions in strains derived from the two case patients and several controls, both local and from the databases, led the authors to conclude that the *pol* sequence on its own did not provide enough information to clarify the relationship between the two patients. This conclusion was questioned [1] on the basis of the use of one single phylogenetic method, neighbour joining on the Kimura two-parameter substitution model, to obtain the corresponding phylogenetic trees. Stürmer *et al.* [2] subsequently verified that their conclusions held after a range of methods was applied to the original data set. As Posada and Crandall [4] have already shown the necessity of applying a correct model of substitution for the inference of relationships by using HIV sequences, we will concentrate here on the original question, that of the feasibility of using one single gene for inferring relationships in transmission cases.

All methods of phylogenetic reconstruction from nucleotide sequences are based on the molecular clock hypothesis [5], according to which mutations occur at random and independently at a roughly constant rate since the divergence of any two sequences from a common ancestor. As there are a limited number of nucleotide positions available for comparison in all cases and there is also a limited number of alternative states (the four bases) at each nucleotide site, there is an upper limit on the number of substitutions that can be detected, which leads to a saturation effect with longer times since divergence. The real amount of divergence can only be inferred from the observed differences between two sequences after the application of an evolutionary model that will correct for superimposed, undetected substitutions. The absolute number of substitutions per site, the basic parameter computed in methods that reconstruct phylogenies from pair-wise distances such as neighbour-joining or minimum evolution, depends on the product of the rate of evolution and the time since divergence. The longer this time, the larger the amount of differences actually produced and the closer we get to the saturation point, at which the linear relationship between divergence time and genetic differentiation is lost [6].

One important feature of HIV-1 is its high evolutionary rate, but this is heterogeneous along its genome [7]. In consequence, the determination of phylogenetic relationships has to be based on the analysis of a region with an adequate level of evolution for the time scale of the problem. Recent transmissions have to be analysed with fast evolving regions (e.g. *env* or *vpu*), because otherwise there would be no or little variation to distinguish two recent events from older events in the same population, whereas older events can only be studied with more slowly evolving genome regions (e.g. *pol*, *gag* or *vif*), as fast evolving regions may have already reached or be close to the saturation point and will no longer provide enough resolution power. This is illustrated in the discordant results obtained by Stürmer *et al.* [3] for the two regions studied with patients R004 and R016. In our interpretation, the analysis of the C2V3 region in the *env* gene did not show significant bootstrap support because enough time had elapsed since the transmission for these two sequences to accumulate almost as many differences as any two control sequences. In contrast to Stürmer *et al.* [3], we conclude that the C2V3 region is inappropriate for resolving the transmission event implicated in patients R004 and R016.

We have developed an alternative methodology for the forensic analysis of transmission events based on nucleotide sequences of RNA viruses [8–10], which involves the study of a number of clones derived from each patient in the appropriate genome region. Although these cases corresponded to hepatitis C virus transmission, the methodology and principles also apply to HIV. The main advantage of this method is that it allows for an individual assessment of the chances of two patients being implied in a transmission case without resorting to the interpretation of bootstrap support values.

Fernando González-Candelas and Andrés Moya,
Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, 46071 Valencia, Spain.

Received: 25 April 2005; accepted: 21 June 2005.

References

1. Jenwitheesuk E, Liu T. **Single phylogenetic reconstruction method is insufficient to clarify relationships between patient isolates in HIV-1 transmission case.** *AIDS* 2005; **19**:743–744.
2. Stürmer M, Preiser W, Gute P, Nisius G, Doerr HW. **Response to 'Single phylogenetic reconstruction method is insufficient to clarify relationships between patient isolates in HIV-1 transmission case' by Jenwitheesuk and Liu.** *AIDS* 2005; **19**:741–743.

3. Stürmer M, Preiser W, Gute P, Nisius G, Doerr HW. **Phylogenetic analysis of HIV-1 transmission: *pol* gene sequences are insufficient to clarify true relationships between patient isolates.** *AIDS* 2004; **18**:2109–2113.
4. Posada D, Crandall KA. **Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1).** *Mol Biol Evol* 2001; **18**:897–906.
5. Kimura M. *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press; 1983.
6. Li WH. *Molecular evolution*. Sunderland, MA: Sinauer; 1997.
7. Muse SV. **Modeling the molecular evolution of HIV sequences.** In: *The evolution of HIV*. Edited by Crandall KA. Baltimore and London: The Johns Hopkins University Press; 1999. pp. 122–152.
8. González-Candelas F, Bracho MA, Moya A. **Molecular epidemiology and forensic genetics: application to a hepatitis C virus transmission event at a hemodialysis unit.** *J Infect Dis* 2003; **187**:352–358.
9. Moya A, Holmes EC, González-Candelas F. **The population genetics and evolutiony epidemiology of RNA viruses.** *Nat Rev Microbiol* 2004; **2**:279–288.
10. Bracho M, Gosalbes MJ, Blasco D, Moya A, González-Candelas F. **Molecular epidemiology of a hepatitis C virus outbreak in a hemodialysis unit.** *J Clin Microbiol* 2005; **43**:2750–2755.

QAD	
Manuscript No.	200022

AIDS
Typeset by Thomson Digital for
Lippincott Williams & Wilkins

Dear Author,

During the preparation of your manuscript for typesetting, some queries have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin as neatly as possible or compile them as a separate list. This form should then be returned with your marked proof/list of corrections to the Production Editor.

QUERIES: to be answered by AUTHOR/EDITOR

AUTHOR: The following queries have arisen during the editing of your manuscript. Please answer the queries by marking the requisite corrections at the appropriate positions in the text.

QUERY NO.	QUERY DETAILS	
	No Qurey	