



ORAL PRESENTATION

Open Access

Mapping of positive selection sites in the HIV genome in the context of RNA and protein structural constraints

Joke Snoeck^{1,2*}, Jacques Fellay^{1,3}, Daniel C Douek⁴, Amalio Telenti¹

From *Frontiers of Retrovirology* 2011

Amsterdam, The Netherlands. 3-5 October 2011

Background

The HIV-1 genome is subject to great variability. On the one hand, there are recognized pressures that target the virus and resulting in escape and adaptation. On the other hand, there is a requirement for sequence conservation because of functional and structural constraints. Mapping the sites of positive selective pressure on the viral genome generates a reference for understanding the limits to viral escape, and can serve as a template for the discovery of sites of genetic conflict with known or unknown host proteins.

Materials and methods

To build a thorough evolutionary, functional and structural map of the HIV-1 genome, full subtype B sequences were obtained from the Los Alamos database. We mapped sites under positive selective pressure (using the SLAC method in HYPHY), amino acid conservation, protein and RNA structure [1], overlapping coding frames, as well as CD8 T cell, CD4 T cell and antibody epitopes (based on the lists from the Los Alamos database website). Fisher exact test was used for univariate analysis, and either logistic regression or binary Firth's penalized-likelihood logistic regression was used in multivariate analysis. Statistical analyses were performed in R version 2.13.0 (<http://www.r-project.org/>).

Results

Globally, 33% of amino acid positions were found to be variable and 12% of the genome was under positive selection. Because interrelated constraining and

diversifying forces shape the viral genome, we included the variables from both classes of pressure in a multivariate model to predict conservation or positive selection: molecular structures (structured RNA and α -helix domains) predicted conservation while CD4 T cell and antibody epitopes were associated with positive selection. Gene-specific analyses showed instances of departure from the genome-wide estimates. For example, CD8 T cell epitopes were generally well conserved, except in *gp120*, where they were enriched for sites under positive selective pressure. Similarly, CD4 T cell epitopes were enriched for sites under positive selection genome-wide, but these epitopes were significantly more conserved in *gag*. Thus, the various constraints and selective pressures do not act evenly across the genome.

Conclusions

Because the global map of the viral genome also identifies positive selected sites that are not in canonical CD8 T cell, CD4 T cell or antibody epitopes, it identifies a class of residues that may be targeted by other host selective pressures, such as innate immunity effectors. As an example, sequence adaptation has been observed in the viral capsid in rhesus macaques upon cross-species transmission of SIVsm, due to selective pressure provided by restrictive TRIMS alleles. In addition to informing the combined analysis of host and viral genetic information, the results of this study may therefore help reveal novel mechanisms of antiretroviral response.

Author details

¹Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland. ²Rega Institute for Medical Research, KU Leuven, Leuven, Belgium. ³Global Health Institute, School of Life Sciences,

¹Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland

Full list of author information is available at the end of the article

EPFL, Lausanne, Switzerland. ⁴Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland, USA.

Published: 3 October 2011

Reference

1. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr, Swanstrom R, Burch CL, Weeks KM: **Architecture and secondary structure of an entire HIV-1 RNA genome.** *Nature* 2009, **460**:711-716.

doi:10.1186/1742-4690-8-S2-O42

Cite this article as: Snoeck *et al.*: Mapping of positive selection sites in the HIV genome in the context of RNA and protein structural constraints. *Retrovirology* 2011 **8**(Suppl 2):O42.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

