## Poster C-39

**Evaluation of Brazilian web based program for classification of HIV-1 sequences**


**Authors:**
Luciano V Araujo (*Department of Computer Science - Department of Computer Science - University of São Paulo*)
João E Ferreira (*Department of Computer Science - University of São Paulo*)
Sabri S Sanabani (*Fundação Pro-Sangue, Hemocentro, São Paulo, Brazil*)
Ester C Sabino (*Fundação Pro-Sangue, Hemocentro, São Paulo, Brazil*)

**Short Abstract:** An automated web based tool for assigning HIV-1 pure and recombinant subtypes within unaligned sequences is presented. The system combines the BLAST search algorithm and the recombination identification program for genetic subtyping of HIV-1. The softer was validated through combined analysis of simulated and other HIV-1 real data.

**Long Abstract:**

BACKGROUND
The widespread use of HIV drug resistance testing results in a phenomenal increase of HIV-1 protease (PR) and reverse transcriptase (RT) sequences. Approximately 5000 to 10.000 genotype tests are performed every year by the Brazilian AIDS program laboratory network. The incorporation of HIV subtype information to such volume of clinical data has created the need for the development of automated computer tool to aid identification of HIV-1 subtypes. Here, we developed an easy to use web-based software that quickly and accurately identifies HIV-1 pure and recombinant subtypes.

DESIGN AND RESULTS
The software is an interface tool that uses the BLAST-based homology search algorithm and the recombination identification program (RIP) to assign HIV-1 subtype prediction. A sequence is subtyped if both programs gave equivalent results, otherwise the sequence is considered untypable. The outcomes of the analysis are presented in sliding window approach. The width of the window is open for both programs while the moving steps and statistical level parameters are only open for the RIP program. Optimal parameters for detection of recombination were determined from 7890 simulated recombinant data performed by insertion of fragments of various sizes ranged from 50-600 bp using a non-reciprocal exchange between a donor (subtype F) and a receptor sequence (subtype B). The performance of the softer was evaluated through combined analysis of simulated and 929 real data from the PR and RT region using different window sizes between 50 bp and 600 bp for both programs. Based on simulated data, the window of 200 bp was consistently revealed clearest signal and therefore chosen as an optimal width parameter to detect recombination. Of the 929 real sequences used to define the subtype alignment, 95.37% were correctly subtyped and only 4.63% were reported untypable that may need more sophisticated phylogenetic analysis tools to assign their subtypes.

CONCLUSION
Our program has greatly facilitated identification of HIV subtypes by allowing batch

submission of unaligned sequences and showed a good performance to be used in the Brazilian AIDS program.