

## **Abstract**

The envelope (env) protein of human immunodeficiency virus type 1 (HIV-1) plays a crucial role in virus entry and is a central target for HIV vaccine design. Using the QUASI program, we analyzed the conserved regions of all currently available env sequences in the Los Alamos National Laboratory HIV Sequence Database and identified positive selection (PS) sites that are likely to be restricted by host immune responses. We found that PS sites are dispersed across conserved regions of env sequence, and that the C3, C4, and C5 regions were the most targeted. Several regions were identified as being PS free and were mainly distributed in the C1 and C2 regions. When comparing individual QUASI PS site frequencies across clades and geographical regions with the overall frequency of the entire env database, the env sequences from North America showed significantly lower PS site frequency, while those from Asia were significantly higher using Student's t test. The QUASI PS site frequency of env proteins from viruses isolated from different years showed that the PS site frequencies of the env population increased over time. Our study provides an overview of PS sites across the conserved regions of HIV-1 env sequences.