

## Figure Legend.

### Figure 1. Evolutionary relationships of the viral sequences in *env* gene

**a)** The evolutionary history was inferred using the Neighbor-Joining method. The tree is drawn to scale. The evolutionary distances were computed using the Maximum Composite Likelihood method. The rate variation among sites was modeled with a gamma distribution (shape parameter = 0.7166). Numbers at branch nodes refer to the bootstrap support (only values greater than 95% are shown). The sequences in *env* gene of the patients considered dual infected (DI 03, DI 06, DI 08, DI 16, DI 52, DI 56 and DI 72) are marked with brown, green, red, blue, pink, turkey and grey circle respectively; **A and B letters identify the viral populations found in DI patients.** **b)** Distribution of the genetic distances between individuals' subtype B virus in *env* gene. The number of base substitutions per site from averaging over all sequence pairs between MI patient's nucleotide sequences was calculated. Analyses were conducted using the Maximum Composite Likelihood model. The rate variation among sites was modeled with a gamma distribution (shape parameter = 0.7166) (MEGA 6). Genetic distance values obtained were divided into 2% intervals, and the number of comparisons in each interval was plotted against the genetic distance. Red line indicated genetic distance between the two viruses in each intra-subtype B DI patient identified. Blue line indicated genetic distance between the two viruses in each inter-subtype DI patient identified that were not included in the estimates.