



Additional file 4: Maximum likelihood phylogenetic analysis of the assembled sequences from the ALIVE cohort with reference sequences representing subtypes A, B, C, D, CRF01_AE and CRF02_AG using the Bayes Information Criterion in MEGA6. Reference sequences are indicated by the subtype following by the Genbank accession number. The reliability of the clustering patterns were tested by bootstrapping and 1000 bootstrap pseudo-samples were used.