ANALYSIS OF DATA

Figure A (PAUP Tree). The tree produced by PAUP* was done first using a heuristic search of about 30,000 trees, and then a bootstrap analysis; the combination of the two led to this tree being outputted as it had the highest bootstrap scores. The numbers on the branches represent the bootstrap values extracted; a value of 90 or above demonstrates incredibly strong support, a quality which nearly all of the branches satisfy.

 Each colored box represents a clade, a group of organisms that share a common ancestor. The big blue box represents the clade that contains all of the HIV-1 sequences, the pale box represents the clade that contains all of the HIV-2 sequences, and other clades are colored for comparison and labeling purposes.

Figure B (BEAST Tree). The tree produced by BEAST was produced using a strict molecular clock, an exponential clock rate, and the MCMC chain set at 20,000,000. The numbers on the branches represent posterior probabilities; as such, values closer to 1 mean better supported branches. This tree has incredibly high posterior probabilities, with all but one branch having probabilities of 0.999 or greater.

The same coloring scheme was used, and a very similar tree was produced with only one or two major differences.

Analysis. The first thing to notice is the raw similarity - it shows that this tree is likely the true phylogenetic tree, as it was replicated almost entirely using two entirely different algorithms. Their high probabilities also increase this likeliness of being the true tree. The main pattern to be noticed is the clear split between HIV-1 and HIV-2: the two take up completely different branches and as a result are quite different. Another pattern to be noted is that the two also identify almost entirely with SIVcpz (chimpanzee SIV) and SIVsmm (sooty mangabey SIV) respectively.