METHODS

In constructing the trees, three main steps were taken, using a variety of programs:

1) Gathering and Alignment of Sequences. Using online databases (primarily hiv.lanl.gov), about 80 different sequences were gathered, with the final trees consisting of 55 sequences. These sequences would then be opened in SeaView, a GUI for viewing sequences, and then aligned using MAFFT and its Fast Fourier Transformations.

2) Creation of Trees. The two main programs used were PAUP* and BEAST for this step:

> **PAUP* (Phylogenetic Analysis Using Parsimony)**. The first program used, PAUP took the aligned sequences and performed **parsimony**-constraint heuristic search and **bootstrap** analysis of over 100 different trees.

> **BEAST (Bayesian Evolutionary Analysis Sampling Trees).** The second program used; an XML file was created using BEAUTi where the **MCMC** chain was specified at 20,000,000 and a strict **molecular clock** was used. Each of the 20,000 trees produced by the BEAST run was weighted based on its posterior probability, and the best tree was annotated using another program called TreeAnnotator. The annotated tree was then viewed in FigTree, and statistical results from the MCMC analysis was viewed in Tracer.

Between the two programs, almost 50 distinct trees were created until one was found with suitable bootstrap and/or posterior probability.

3) Revision of Gene Pool and Outgroups. To produce high-quality trees, many deletions and additions were made to the sequences used; each combination was then tested to see how it performed in PAUP. Almost 40 different changes were tested before creating the final tree. PAUP* also requires the user to set an **outgroup**, a species known to have diverged before the species in study. As a result of this, about 15 different outgroups were attempted before finding one with both good scientific and phylogenetic backing.