

# A Phylogenetic Analysis of the HIV-1 and HIV-2 Strains

## Initial Research Plan

---

### Introduction / Rationale for Doing Project

The **human immunodeficiency virus**, or **HIV**, is a relatively new pandemic, responsible for causing AIDS and killing millions of people around the world. It is divided into 2 main types, HIV-1 and HIV-2. More well-known is the former; it is responsible for the vast majority of AIDS cases and has spread worldwide while HIV-2 is prevalent mainly in West Africa and is less pathogenic.

In order to track evolutionary relationships and rates/dates of divergence, scientists have created phylogenetic trees with both subtypes of HIV-1 and 2. In doing so, scientists have discovered that HIV evolved from **simian immunodeficiency virus**, or **SIV**, a closely related disease found in non-human primates.

However, many of the trees are of either low or unpublished bootstrap and/or posterior probability, both of which are indicators of a strong tree. Much focus has also been put on tracking the subtypes of HIV-1 and 2 more common in the Western world, and not so much in the rest of the world. As a result, we will attempt to create a well-supported tree with HIV-1, HIV-2 and SIV sequences from West Africa, an area of less study but an area with a large diversity of HIV subtypes.

---

### Hypothesis

By inspecting West African HIV-1, HIV-2 and SIV genomes, it is hypothesized that information will be produced that ties together the three viruses in a well-supported phylogenetic tree. From the tree, it is expected that we will find evidence for relationships found in different regions as well as discovering some of our own new theories regarding the evolution of the virus.

---

### Materials / Procedure

#### Materials

1. Modern computer with proper software downloaded and access to internet
2. A Nucleotide Database: [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov) or [hiv.lanl.gov](http://hiv.lanl.gov)
3. Software
  - a. Download *KomodoEdit* (or any plain text editor), *Seaview*, *Mafft*, *Mesquite*, *BEAST*, *BEAUTi*, *Tree Annotator*, *Tracer*, *FigTree*, *PAUP\**

#### Procedure

1. Acquire all programs listed in “Materials”.
2. Acquire data
  - a. Go to different nucleotide databases (e.g. [hiv.lanl.gov](http://hiv.lanl.gov), a database for HIV) and search for the necessary sequences.

- b. Export a .fasta file with nucleotide sequences.
  3. Aligning the File with *Seaview* / *Mafft*
    - a. Open the .fasta file in *Seaview*
    - b. Using the built-in alignment programs or *Mafft*, align the sequence
    - c. Optionally, you can manually adjust alignments with *Mesquite*
    - d. Save aligned sequences in NEXUS format (.nex)
  4. Create Tree in PAUP\*
    - a. Run Heuristic Search
    - b. Run Bootstrap Analysis
  5. Setting the Parameters for *BEAST*
    - a. Open the .nex file in *BEAUTi*
    - b. Separate the taxa into taxon sets depending on which sets are of interest and are related to one another
    - c. Set the dates of each taxa through “set tip dates” and setting them by hand if needed
    - d. Click on Clock Models and change the Molecular Clock Model to a Strict Clock Model
    - e. Leave all else at the default settings unless you are willing to explore the other options that *BEAUTi* offers.
    - f. Click “Generate BEAST File” and an .xml file will be created
  6. Input the .xml file into *BEAST* and click “Run”.
    - a. When *BEAST* is finished, it will produce a log file and a tree file in the same location as your .xml file
  7. Open the log file in *Tracer*
    - a. *Tracer* provides various graphs and quantities that can be used to explore *BEAST*’s output
  8. Open the tree file in *TreeAnnotator*
    - a. *TreeAnnotator* can be used to modify and annotate the tree file
  9. Open the tree file in *FigTree*
    - a. *FigTree* can view, color and change the trees as the user wishes
- 

## Data Collection and Analysis

After collecting the sequences and the trees, they would be analyzed by examining the bootstrap and posterior probability values generated, both of which determine the strength of a particular tree. If the tree is deemed to have sufficient support, we will try to make conclusions.

By inspecting the subtypes of HIV-1, HIV-2, SIV and their relationships within West Africa, we will attempt to find similarities with trees from other regions of the world. We will also attempt to discover new and undiscovered patterns; once found, we will then try to apply this knowledge to potential applications, like treatments for the virus or discovery of new subtypes.

---

## Bibliography

- Baum, David. "Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups." *Scitable*. NatureEducation, 2008. Web. 12 Oct. 2015.  
<<http://www.nature.com/scitable/topicpage/reading-a-phylogenetic-tree-the-meaning-of-41956>>. This article specializes in the reading and usage of an evolutionary tree, providing many illustrations to help detail the author's points. It goes over the different representations of trees, and shows how although they look very different, they portray the same information.
- Baum, David A., and Susan Offner. "Phylogenies and Tree Thinking." *American Biology Teacher* 70.4: 222-29. Print. This source provides an in-depth introduction to the concepts of phylogeny, going over basic definitions and terms while providing detailed examples with each topic. It is fairly accessible and goes into detail on mistakes that students make whilst analyzing trees, and how their use is more and more important in classrooms.
- "Division of AIDS (DAIDS)." *National Institute of Allergy and Infectious Diseases*. Web. 12 Oct. 2015.  
<<http://www.niaid.nih.gov/about/organization/daids/Pages/default.aspx>>. This article provides basic information on the HIV virus, the HIV/AIDS epidemic, and what has been done so far to fight against it.
- "Human Immunodeficiency Virus Type 2 (HIV-2)." *HIV Clinical Resource*. Web. 15 Nov. 2015.  
<<http://www.hivguidelines.org/clinical-guidelines/adults/human-immunodeficiency-virus-type-2-hiv-2/>>. This website provides an introduction to the HIV-2 subtype and talks about its history and effect on the human body. It also briefly discusses differences between it and the HIV-1 subtype as well as treatments for both viruses.
- Nyamweya, Samuel, et al. "Comparing HIV-1 and HIV-2 Infection: Lessons for Viral Immunopathogenesis." *Wiley Online Library*. Web. 15 Nov. 2015.  
<<http://www.ncbi.nlm.nih.gov/pubmed/23444290>>. This article focuses on the differences between the HIV-1 and HIV-2 subtypes. It discusses differences in region, biochemistry, origin, and progression to AIDS. It also discusses how these differences could be used to better current treatment for HIV.
- Thanukos, Anna. "Phylogenetic Systematics, a.k.a. Evolutionary Trees." *Understanding Evolution*. U of California Museum of Paleontology, 2006. Web. 12 Oct. 2015.  
<[http://evolution.berkeley.edu/evolibrary/article/phylogenetics\\_01](http://evolution.berkeley.edu/evolibrary/article/phylogenetics_01)>. This website provides further introductions into evolutionary trees, and defines various terms needed in the field. It also discusses the uses and applications of trees in modern-day biology.