

## **PhyloGenie quick user guide**

PhyloGenie is a generic web framework for phylogenetic inference with GPU compatibility. It supports molecular data types, both Amino acid and DNA. The main objective of the system is to enhance the performance of the phylogenetic inference process by utilizing GPU architecture and visualize the generated phylogenetic tree with multiple visualization formats with better user experience.

Tool comprise of all the phases of phylogenetic Inference process.

1. Data Pre-Processing
2. Multiple Sequence Alignment
3. Algorithm Recommendation
4. Tree Generation
5. Tree Visualization

PhyloGenie supports divergent GPU compatible inference algorithms both character based and distance based.

- Character based
  1. Maximum Likelihood
  2. Maximum Parsimony
  3. Bayesian
- Distance based
  1. Neighbor-Joining
  2. UPGMA

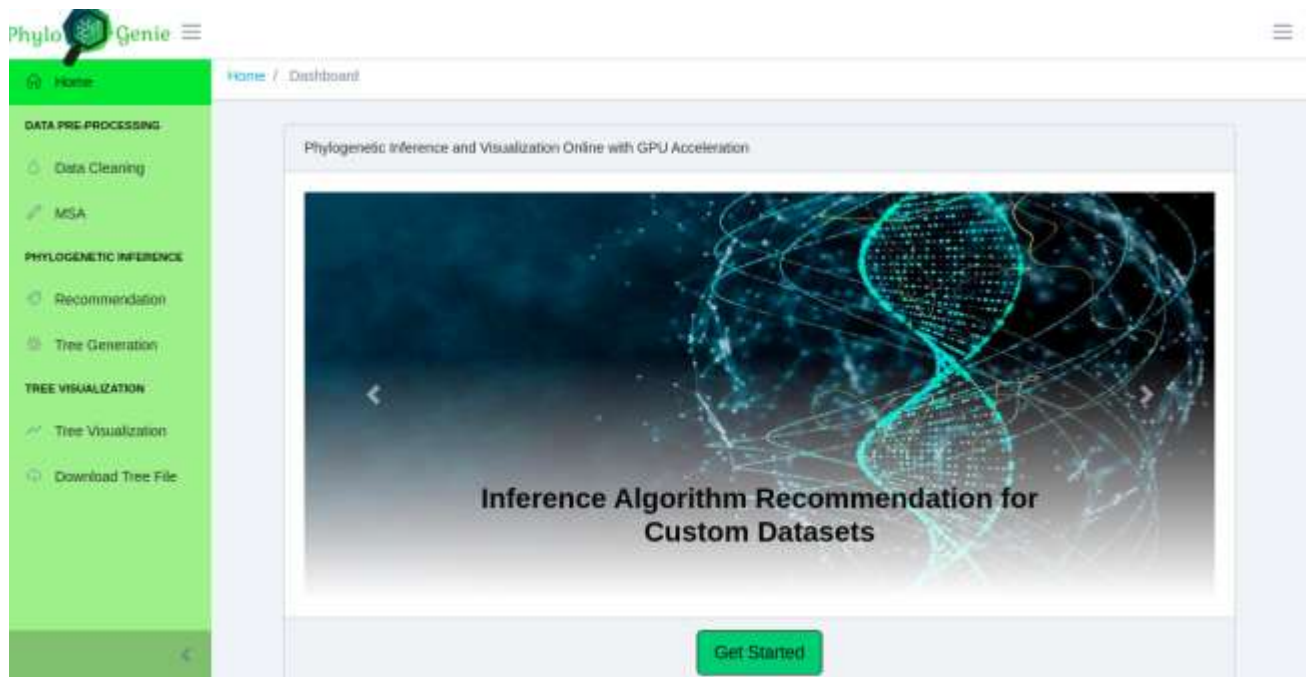
Furthermore tool improves the system usability with algorithm recommendation and providing more interactive visualization techniques. This document will help you to navigate through the PhyloGenie website.

**Website url -**

### **Input Data File**

The user will be provided with both DNA and Amino Acid data sets of same sequence length with different taxa sizes. For each data type there will be data files with 25, 50, 100, 200, 400, 600, 800 and 1000 number of sequences. All the data files will be in FASTA format.

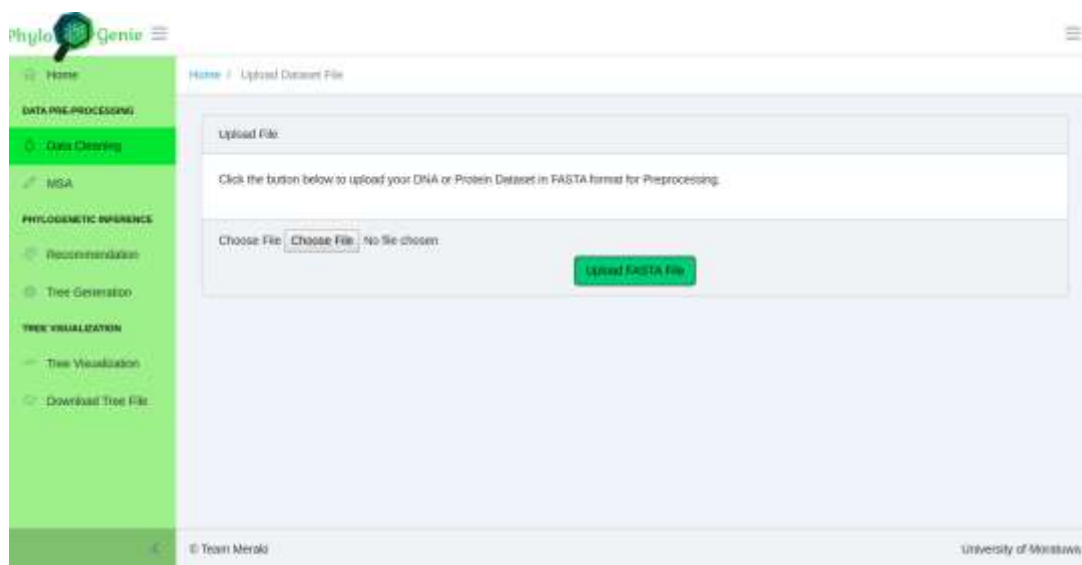
## Main GUI of the System



## Task List

1. Upload your raw data file in FASTA format.

A preview of the uploaded data file will be shown in the right side and after confirming the correct file click upload Fasta File button to proceed.



2. Then data file will be preprocessed and undergo multiple sequence alignment. User can download the aligned data file if required.
3. Tool itself provides five algorithms to generate the phylogenetic tree (Maximum Likelihood, Maximum Parsimony, Bayesian, Neighbor-Joining and UPGMA). And the best fit algorithm for the uploaded dataset file will be recommended and listed with all other available algorithms. User can either select the recommended algorithm or any other preferred algorithm if required.

This screenshot shows a web interface section titled "Recommended Algorithms:". It contains two groups of radio buttons. The first group, under the heading "Recommended Algorithms:", has two options: "Bayesian" (which is selected) and "Maximum Likelihood". The second group, under the heading "Other Available Algorithms", has three options: "NJ", "UPGMA", and "Maximum Parsimony". Below these options is a green button labeled "Infer Tree".

4. Then click infer tree button. Then the system will analyze the features of aligned data file such as the size and its type and based on the system specified threshold values it will select either GPU version or serial version of the recommended algorithm by considering performance factors. User can experience a better performance with less execution time for generating the phylogenetic tree with these GPU compatible algorithms in PhyloGenie with respect to other existing inference tools.

This screenshot shows the full PhyloGenie web interface. On the left is a green sidebar with navigation links: Home, DATA PRE-PROCESSING (with sub-links Data Cleaning, MSA), PHYLOGENETIC INFERENCE (with sub-links Recommendation, Tree Generation), TREE VISUALIZATION (with sub-links Tree Visualization, Download Tree File), and a footer with a copyright notice for Team Meraki. The main content area shows a workflow: "Choose File" (with a "Choose File" button and "01\_n25.fas" text), followed by an "Upload FASTA File" button. Below this is a status message "Dataset preprocessed and aligned successfully" and a "Download Aligned Dataset" button. The "Recommended Algorithms:" section is visible, showing "Bayesian" as the recommended algorithm. At the bottom of this section is an "Infer Tree" button. The footer of the main area includes "© Team Meraki" and "University of Moratuwa".

- Next the generated phylogenetic tree will be displayed in newick format as shown below.

**Recommended Algorithms:**

- ☒ Bayesian
- ☐ Maximum Likelihood

**Other Available Algorithms**

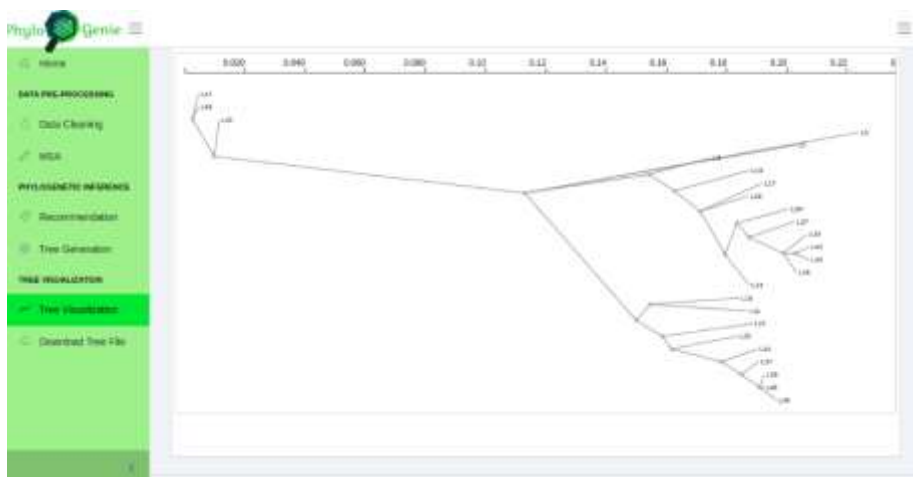
- ☐ NJ
- ☐ UPGMA
- ☐ Maximum Parsimony

**Infer Tree**

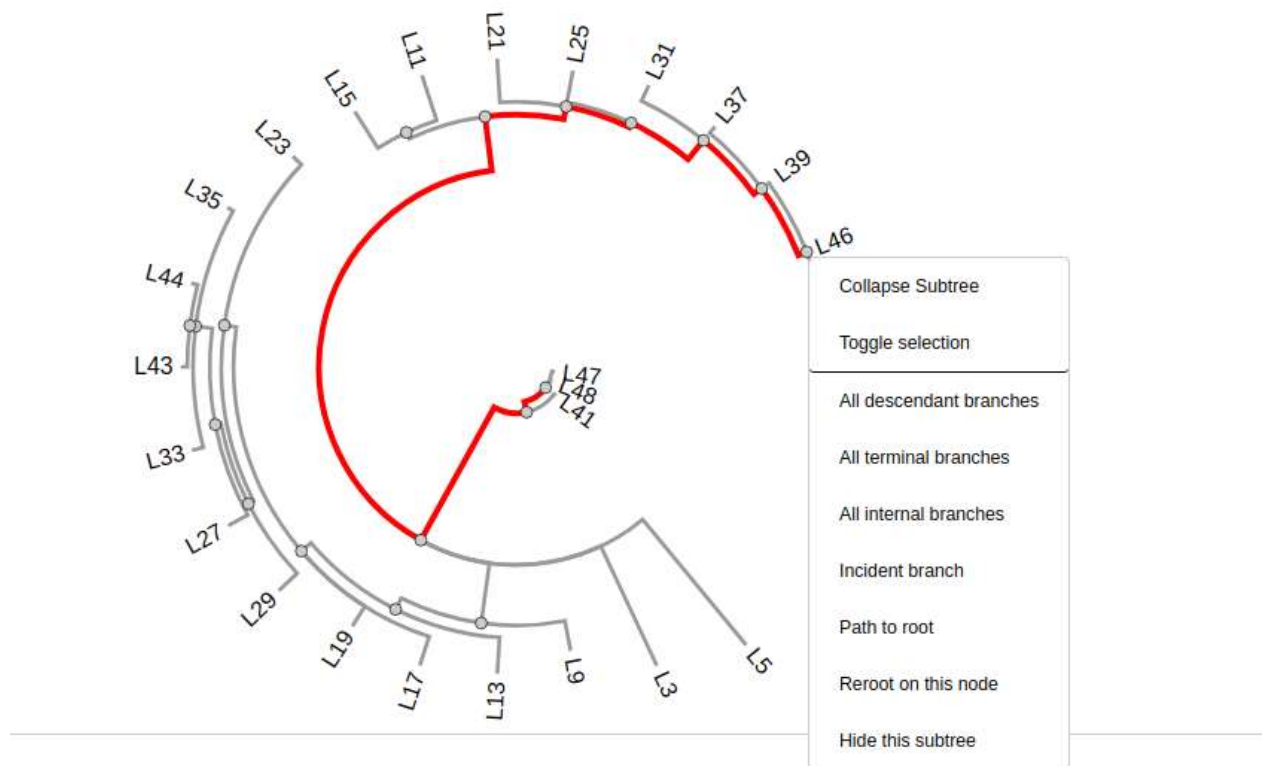
```
(L47:0.00257,L48:0.00315,(L41:0.00289,((L5:0.14377,((L15:0.04297,L11:0.04524):0.00532,(L21:0.03814,(L25:0.02910,(L31:0.01295,(L37:0.00852,(L39:0.00196,L46:0.00292,L45:0.00904):0.00885):0.00748):0.02080):0.00357):0.00937):0.04830):0.00592,L3:0.10954,(L9:0.02686,(L13:0.03255,(L17:0.02847,L19:0.02042,((L29:0.02123,(L27:0.02069,(L33:0.01097,(L43:0.00513,L44:0.00490):0.00500,L35:0.00589):0.01301):0.00546):0.00514,L23:0.01168):0.01002):0.00887):0.01128):0.05243):0.13561):0.00967):0.00000;
```

**Visualize Tree**

- Then the tree can be visualized. PhyloGenie consist of two basic tree visualization formats as cladogram and phylogram in all tree patterns such as rectangular, circular and radial.



7. PhyloGenie provides a more interactive visualization experience with a set of well-defined functionalities as shown below.



8. As final activity user can export the generated Newick file of the inferred phylogenetic tree to their google drive or to local file system.
9. As an external feature PhyloGenie supports visualization of any existing Newick file. Any external tree file can be imported to the website or user can copy paste their tree in Newick format in the provided panel in order to proceed with visualization.

