

PhyloSuite: Suite of tools for streamlining the analysis, manipulation and visualization of phylogenetic data

Jason Charamis^{*1,2}, Pavlos Pavlidis^{1,3}



¹Department of Biology, University of Crete, Vassilika Vouton, Heraklion, Greece.

²Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Heraklion, Greece.

³Institute of Computer Science, Foundation for Research and Technology-Hellas, Heraklion, Greece.

^{*}Corresponding author: jason_charamis@imbb.forth.gr

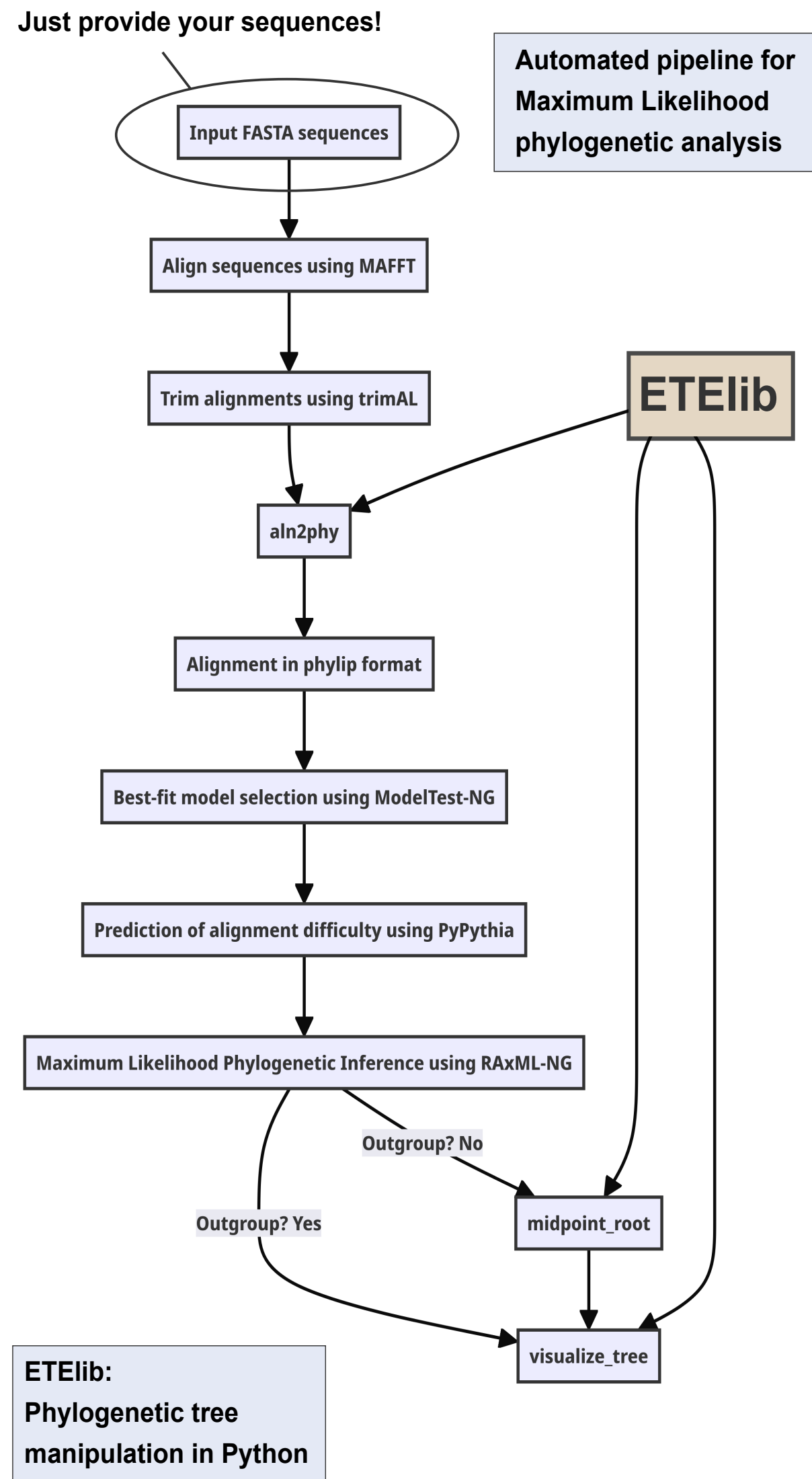


ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE

Abstract

Effective analysis, manipulation and visualization of phylogenetic data is of great importance in any Comparative and Evolutionary Genomics study. While a plethora of tools for these processes currently exist, they often lack integration into easy-to-use workflows. Here we present a suite of three complementary tools, RAxML-NG-and-Friends, ETElib and OliveTRee, which have been designed to streamline the analysis, manipulation and visualization of phylogenetic data with minimal effort by the user. RAxML-NG-and-Friends is an automated Snakemake pipeline for performing maximum likelihood phylogenetic analysis. ETElib is a Python library based on the ETE3 toolkit, with a number of useful functions for effective tree manipulation and basic tree visualization capabilities. OliveTRee is an R package for advanced tree manipulation and visualization. PhyloSuite has been designed to provide a unified framework of analyzing, manipulating and visualizing phylogenetic data at scale.

RAxML-NG-and-Friends



OliveTRee

Advanced visualization of large-scale phylogenies with a few lines of code

Demonstrating OliveTRee's capabilities on the *Squamata* phylogenomic tree with 6,885 species.
Title & Singhal *et al.* (2024). The macroevolutionary singularity of snakes. *Science*. 383(6685), 918–923.

Visualize the *Squamata* phylogenomic tree (n=6,885 taxa) with snake species highlighted

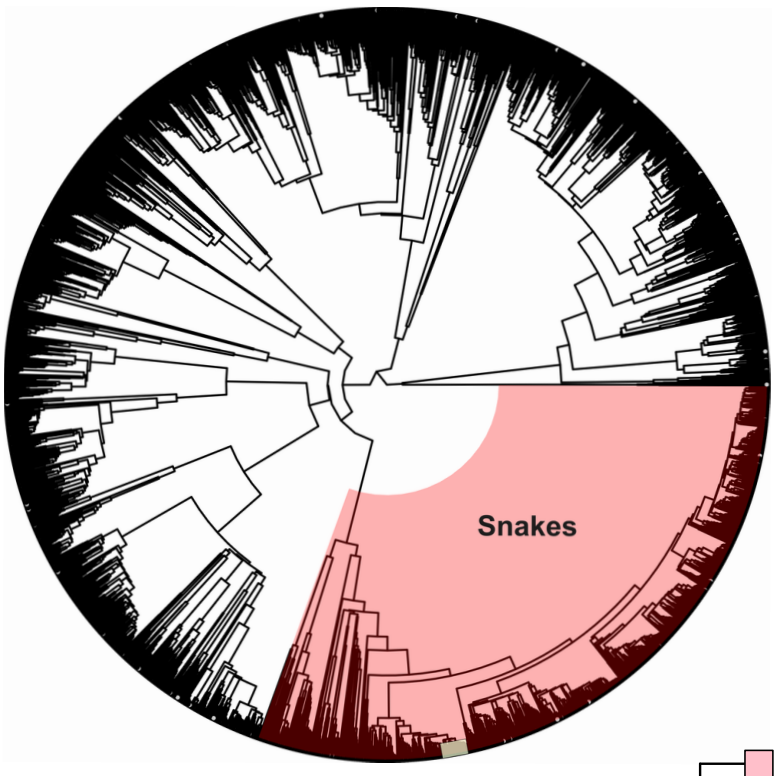
```
source("~/bin/PhyloSuite/OliveTRee/R/OliveTRee.R")
setwd("~/Documents/OliveTRees/R")

# First read the whole tree with 6,885 taxa
squamate_tree <- read_tree(input_file = "squamates_Title_Science2024_ultrametric_constrained.tre")

# Identify parent node of Snakes
print_internal_nodes(squamate_tree, form = "circular")

highlight_nodes = c(
  "Snakes" = 6890
)

# Generate the highlighted tree with colors corresponding to the different clans
highlight_tree(
  squamate_tree,
  form = "circular",
  highlight_nodes = highlight_nodes,
  colors = c("Snakes" = "red"),
  legend_orientation = "vertical",
  legend_key_size = 0.8,
  legend_font_face = "plain",
  legend_spacing_y = 1,
  legend_key_width = 0.8,
)
```



Extract and visualize *Crotalus* subtree (n=51 taxa), with color and shape mappings for each species

```
source("~/bin/PhyloSuite/OliveTRee/R/OliveTRee.R")
setwd("~/Documents/PhyloSuite/R")

# Load file with full species taxonomy
squamates <- read.csv("squamateTreetaxa.csv")

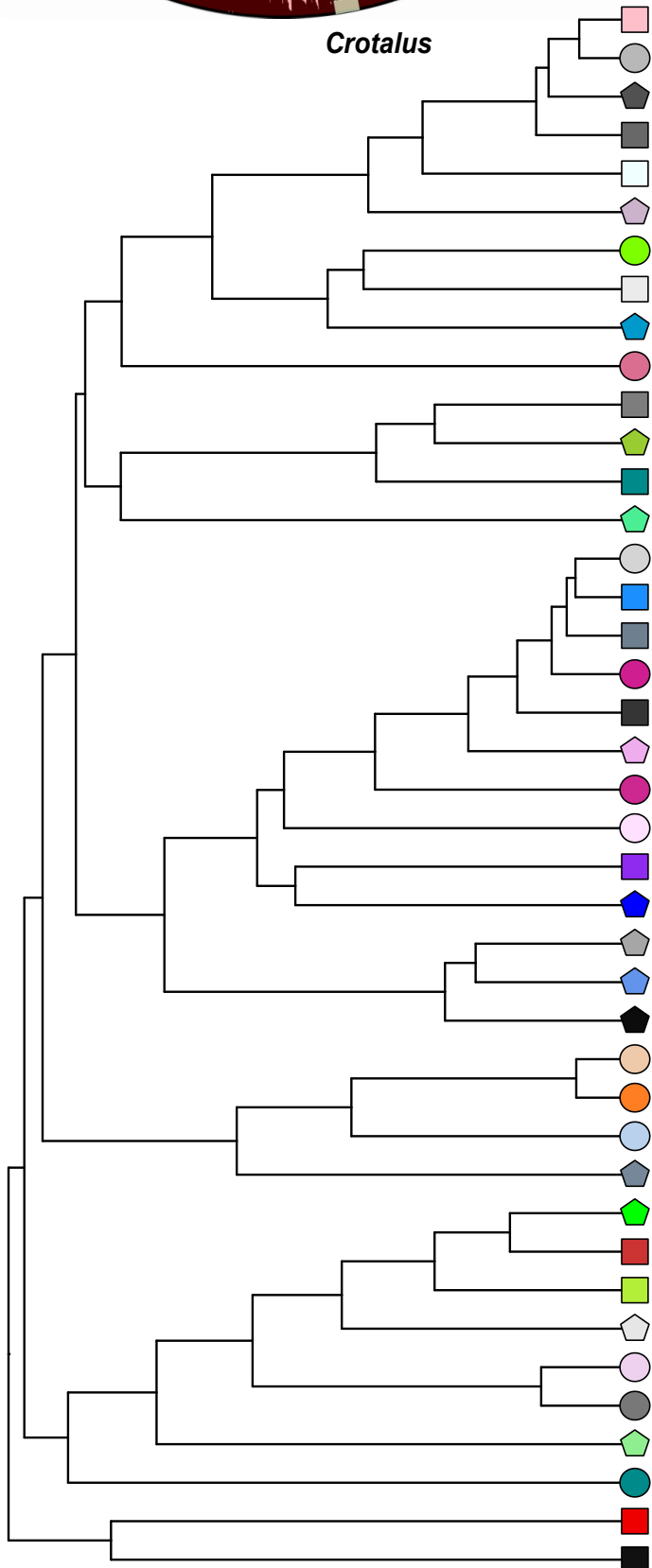
# Extract Crotalus genus subtree
crotalus_species <- c(squamates[grepl('Crotalus', squamates[, 'genus']), ], 1)
crotalus_subtree <- extract_subtree(tree = squamate_tree, taxon_list = crotalus_species)

# Generate random colors for each species
group_colors <- setNames(sample(colors(), length(crotalus_subtree$tip.label)), crotalus_subtree$tip.label)

# Generate random shapes for each species from 'circle', 'triangle' and 'regular pentagon'
group_shapes <- setNames(
  sample(
    c("circle", "square", "regular pentagon"),
    length(crotalus_subtree$tip.label), replace = TRUE
  ),
  crotalus_subtree$tip.label
)

visualize_tree(
  crotalus_subtree,
  color = group_colors,
  shape = group_shapes,
  form = "rectangular",
  legend_position = "right",
  legend_text_position = "right",
  legend_orientation = "vertical",
  legend_key_size = 0.000001,
  legend_font = "Arial",
  legend_fontsize = 5,
  legend_font_face = "italic",
  legend_spacing_x = 0.0000001,
  legend_spacing_y = 0.0000001,
  legend_key_width = 0.00000001,
  legend_title_hjust = 0.00000001,
  legend_name = "Species",
  tiplabels = FALSE,
  taxon_group_separator = NULL,
  tip_label_color = NULL,
  tip_shape_size = 5,
  branch_length = TRUE
)
```

Crotalus simus	Crotalus lutosus	Crotalus tancitarensis
Crotalus unicolor	Crotalus concolor	Crotalus intermedius
Crotalus durissus	Crotalus helleri	Crotalus pricei
Crotalus vegrandis	Crotalus oreganus	Crotalus morulus
Crotalus tzabcan	Crotalus cerberus	Crotalus aquilus
Crotalus culminatus	Crotalus viridis	Crotalus triseriatus
Crotalus molossus	Crotalus scutulatus	Crotalus lepidus
Crotalus totonacus	Crotalus tigris	Crotalus pusillus
Crotalus basiliscus	Crotalus mitchellii	Crotalus armstrongi
Crotalus willardi	Crotalus adamanteus	Crotalus ravus
Crotalus ericsmithi	Crotalus atrox	Crotalus polystictus
Crotalus lannomi	Crotalus catalinensis	Crotalus enyo
Crotalus stejnegeri	Crotalus ruber	Crotalus cerastes
Crotalus horridus	Crotalus transversus	



Conclusion

PhyloSuite facilitates the analysis, manipulation, and visualization of large-scale phylogenetic data. It requires minimal effort by the user, while also vastly reducing the amount of manual work typically needed to perform such processes. Last, PhyloSuite contributes towards reproducibility in phylogenetics.