Distributions were compared using a Kolmogorov-Smirnov test (KS) on the empirical cumulative distribution functions. This test was performed using R’s built-in implementation, ks.test(). Where distributions contained an inequal number of samples, the distribution with more samples was systematically down sampled by retaining samples evenly spaced through the observed distribution.

To represent substitution distributions of subtribes, the brassica ingroup, the non-brassica outgroup, and the last common ancestor of the included species, simulated distributions were generated. Distributions were simulated by drawing samples from each species’ substitution distribution for those species in a given set (e.g., the members of a subtribe). The samples drawn from each species is dictated by the total number of samples times the desired proportion in the simulated distribution rounded to the nearest integer. The total number of samples was defined as the average number of samples among the species to be used, rounded to the nearest integer. The desired proportion for each species was determined using the relatedness in the phylogenetic tree. For each species the path length to every parent node was calculated and used to find the distance from each node to its child species. The representation of each species’ distribution at a given node is assumed to be inversely proportionate to the distance between that node and the species. In cases where one or more species is removed from consideration (e.g., simulating the distribution of the non-brassica outgroup) the expected fractions of the disallowed species is set to zero and then the fractions are normalized to sum to one. Although this approach results in stochastic results – samples are randomly drawn – the sheer number of samples being drawn (17,223 – 167,646 samples) reduces the expected variation between simulations.

Figures associated with this analysis were generated with R 4.2.2 “Innocent and Trusting” #CITE# with the use of tidyverse #CITE# for data manipulation, organization, and plotting, cowplot #CITE# and mdthemes #CITE# for graph arrangement and aesthetic alterations and ape (version 5.6-2) #CITE# and tidytree #CITE# for phylogenetic tree manipulation and tictoc #CITE# for code timing during development. Custom R functions were created for simulating substitution distributions based on phylogenetic tree and distributions for each species and for comparison and visualization of empirical cumulative distribution functions. For details please refer to the associated repository on [GitHub](http://www.github.com/danielkick/brassica_ks) or Zenodo #DOI TBD#.

To cite R in publications use:

R Core Team (2022). R: A language and environment for

statistical computing. R Foundation for Statistical

Computing, Vienna, Austria. URL https://www.R-project.org/.

@Manual{,

title = {R: A Language and Environment for Statistical Computing},

author = {{R Core Team}},

organization = {R Foundation for Statistical Computing},

address = {Vienna, Austria},

year = {2022},

url = {https://www.R-project.org/},

}

To cite package ‘tidyverse’ in publications use:

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François

R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen

TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel

DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani

H (2019). “Welcome to the tidyverse.” \_Journal of Open

Source Software\_, \*4\*(43), 1686. doi:10.21105/joss.01686

<https://doi.org/10.21105/joss.01686>.

@Article{,

title = {Welcome to the {tidyverse}},

author = {Hadley Wickham and Mara Averick and Jennifer Bryan and Winston Chang and Lucy D'Agostino McGowan and Romain François and Garrett Grolemund and Alex Hayes and Lionel Henry and Jim Hester and Max Kuhn and Thomas Lin Pedersen and Evan Miller and Stephan Milton Bache and Kirill Müller and Jeroen Ooms and David Robinson and Dana Paige Seidel and Vitalie Spinu and Kohske Takahashi and Davis Vaughan and Claus Wilke and Kara Woo and Hiroaki Yutani},

year = {2019},

journal = {Journal of Open Source Software},

volume = {4},

number = {43},

pages = {1686},

doi = {10.21105/joss.01686},

}

To cite package ‘cowplot’ in publications use:

Wilke C (2020). \_cowplot: Streamlined Plot Theme and

Plot Annotations for 'ggplot2'\_. R package version

1.1.1, <https://CRAN.R-project.org/package=cowplot>.

@Manual{,

title = {cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'},

author = {Claus O. Wilke},

year = {2020},

note = {R package version 1.1.1},

url = {https://CRAN.R-project.org/package=cowplot},

}

To cite package ‘mdthemes’ in publications use:

Neitmann T (2020). \_mdthemes: Markdown Themes for

'ggplot2'\_. R package version 0.1.0,

<https://CRAN.R-project.org/package=mdthemes>.

@Manual{,

title = {mdthemes: Markdown Themes for 'ggplot2'},

author = {Thomas Neitmann},

year = {2020},

note = {R package version 0.1.0},

url = {https://CRAN.R-project.org/package=mdthemes},

}

To cite package ‘tictoc’ in publications use:

Izrailev S (2022). \_tictoc: Functions for Timing R Scripts,

as Well as Implementations of "Stack" and "List"

Structures\_. R package version 1.1,

<https://CRAN.R-project.org/package=tictoc>.

A BibTeX entry for LaTeX users is

@Manual{,

title = {tictoc: Functions for Timing R Scripts, as Well as Implementations of

"Stack" and "List" Structures},

author = {Sergei Izrailev},

year = {2022},

note = {R package version 1.1},

url = {https://CRAN.R-project.org/package=tictoc},

}

ATTENTION: This citation information has been auto-generated

from the package DESCRIPTION file and may need manual editing,

see ‘help("citation")’.

To cite ape in a publication please use:

Paradis E. & Schliep K. 2019. ape 5.0: an environment for

modern phylogenetics and evolutionary analyses in R.

Bioinformatics 35: 526-528.

@Article{,

title = {ape 5.0: an environment for modern phylogenetics and evolutionary analyses in {R}},

author = {E. Paradis and K. Schliep},

journal = {Bioinformatics},

year = {2019},

volume = {35},

pages = {526-528},

}

As ape is evolving quickly, you may want to cite also its

version number (found with 'library(help = ape)' or

'packageVersion("ape")').

To cite tidytree in publications use:

Guangchuang Yu. (2022). Data Integration, Manipulation and

Visualization of Phylogenetic Trees (1st edition). Chapman

and Hall/CRC. doi:10.1201/9781003279242

@Book{,

title = {Data Integration, Manipulation and Visualization of Phylogenetic Treess},

author = {Guangchuang Yu},

publisher = {Chapman and Hall/{CRC}},

year = {2022},

edition = {1st edition},

doi = {10.1201/9781003279242},

url = {https://www.amazon.com/Integration-Manipulation-Visualization-Phylogenetic-Computational-ebook/dp/B0B5NLZR1Z/},

}