

datelife_class

Resources:

collaborative document: <https://pad.carpentries.org/datelife> datelife documentation: <http://phylotastic.org/datelife/index.html>

Installing datelife

```
devtools::install_github("phylotastic/datelife")
```

```
## Skipping install of 'datelife' from a github remote, the SHA1 (0c39cee3) has not changed since last  
## Use `force = TRUE` to force installation
```

Get a list of taxonomic names that you are interested on, it can/t be common names

- We can get tip labels from a tree
- We can create a list of species or higher group names
- to standardize to the taxonomy, we use `make_datelife_query`

Case when we have a single taxon name

```
my_group <- "fringillidae"
```

Case when we want names from a tree

```
my_tree <- ape::read.tree(file = "data/krenek_etal_2013.tre")  
my_tree
```

```
##  
## Phylogenetic tree with 97 tips and 96 internal nodes.  
##  
## Tip labels:  
## 'Natronomonas pharaonis DnaK', 'Haloferax volcanii DnaK', 'Methanosarcina barkeri DnaK', 'Methanosarcina  
##  
## Rooted; includes branch lengths.
```

```
str(my_tree)
```

```
## List of 4  
## $ edge : int [1:192, 1:2] 98 99 100 100 99 101 101 102 102 103 ...  
## $ edge.length: num [1:192] 0.0697 0.4477 0.1343 0.1035 0.0676 ...  
## $ Nnode : int 96  
## $ tip.label : chr [1:97] "'Natronomonas pharaonis DnaK'" "'Haloferax volcanii DnaK'" "'Methanosarcina barkeri DnaK'"  
## - attr(*, "class")= chr "phylo"  
## - attr(*, "order")= chr "cladewise"
```

```
my_tree$tip.label
```

```
## [1] "'Natronomonas pharaonis DnaK'"  
## [2] "'Haloferax volcanii DnaK'"  
## [3] "'Methanosarcina barkeri DnaK'"  
## [4] "'Methanosaeta thermophila DnaK'"
```

```

## [5] "'Nitrosopumilus maritimus DnaK'"
## [6] "'Aciduliprofundum boonei DnaK'"
## [7] "'Thermoplasma volcanium DnaK'"
## [8] "'Caenorhabditis elegans BiP2'"
## [9] "'Caenorhabditis elegans BiP1'"
## [10] "'Daphnia pulex BiP'"
## [11] "'Drosophila melanogaster BiP'"
## [12] "'Danio rerio HSPA5'"
## [13] "'Homo sapiens HSPA5'"
## [14] "'Mus musculus BiP'"
## [15] "'Gallus gallus BiP'"
## [16] "'Oxytricha trifallax BiP'"
## [17] "'Ichthyophthirius multifiliis BiP'"
## [18] "'Tetrahymena borealis BiP'"
## [19] "'Tetrahymena malaccensis BiP'"
## [20] "'Tetrahymena thermophila BiP'"
## [21] "'Paramecium caudatum PcHsp70ER1a/b'"
## [22] "'Paramecium tetraurelia Hsp70Pt07'"
## [23] "'Paramecium tetraurelia Hsp70Pt08'"
## [24] "'Paramecium tetraurelia Hsp70Pt09'"
## [25] "'Paramecium caudatum PcHsp70ER2a/b/c'"
## [26] "'Caenorhabditis elegans Hsp70-7'"
## [27] "'Caenorhabditis elegans Hsp70-9'"
## [28] "'Drosophila melanogaster Hsc70-2'"
## [29] "'Drosophila melanogaster Hsp70Ba'"
## [30] "'Drosophila melanogaster Hsp70Aa'"
## [31] "'Caenorhabditis elegans Hsp70-1'"
## [32] "'Drosophila melanogaster Hsc70-1'"
## [33] "'Drosophila melanogaster Hsc70-4'"
## [34] "'Danio rerio HSPA8'"
## [35] "'Homo sapiens HSPA8'"
## [36] "'Mus musculus HSPA8'"
## [37] "'Gallus gallus HSPA8Aa'"
## [38] "'Homo sapiens HSPA6'"
## [39] "'Mus musculus HSPA1A'"
## [40] "'Homo sapiens HSPA1A'"
## [41] "'Mus musculus HSPA1L'"
## [42] "'Homo sapiens HSPA1L'"
## [43] "'Danio rerio HSPA1B'"
## [44] "'Danio rerio HSPA1A'"
## [45] "'Homo sapiens HSPA2'"
## [46] "'Mus musculus HSPA2'"
## [47] "'Gallus gallus HSPA2'"
## [48] "'Daphnia pulex Hsp70-3'"
## [49] "'Daphnia pulex Hsp70-2'"
## [50] "'Daphnia pulex Hsp70-1'"
## [51] "'Oxytricha trifallax hsp70'"
## [52] "'Paramecium caudatum PcHsp70CY2b'"
## [53] "'Paramecium caudatum PcHsp70CY2a'"
## [54] "'Paramecium tetraurelia Hsp70Pt04'"
## [55] "'Paramecium tetraurelia Hsp70Pt03'"
## [56] "'Paramecium tetraurelia Hsp70Pt05'"
## [57] "'Paramecium caudatum PcHsp70CY1a/b'"
## [58] "'Paramecium caudatum PcHsp70CY1c'"

```

```

## [59] "'Paramecium tetraurelia Hsp70Pt01'"
## [60] "'Paramecium tetraurelia Hsp70Pt02'"
## [61] "'Ichthyophthirius multifiliis hsp70'"
## [62] "'Tetrahymena borealis hsp70-4'"
## [63] "'Tetrahymena thermophila hsp70-4'"
## [64] "'Tetrahymena malaccensis hsp70-4'"
## [65] "'Tetrahymena thermophila hsp70-1'"
## [66] "'Tetrahymena malaccensis hsp70-1'"
## [67] "'Tetrahymena borealis hsp70-1'"
## [68] "'Tetrahymena thermophila hsp70-3'"
## [69] "'Tetrahymena malaccensis hsp70-3'"
## [70] "'Tetrahymena borealis hsp70-3'"
## [71] "'Tetrahymena borealis hsp70-5'"
## [72] "'Tetrahymena thermophila hsp70-5'"
## [73] "'Tetrahymena malaccensis hsp70-5'"
## [74] "'Tetrahymena malaccensis hsp70-2'"
## [75] "'Tetrahymena thermophila hsp70-2'"
## [76] "'Oxytricha trifallax mtHsp70'"
## [77] "'Paramecium caudatum PcHsp70MT1a'"
## [78] "'Paramecium tetraurelia mtHsp70-2'"
## [79] "'Paramecium tetraurelia mtHsp70-1'"
## [80] "'Ichthyophthirius multifiliis mtHsp70-1'"
## [81] "'Tetrahymena thermophila mtHsp70'"
## [82] "'Tetrahymena malaccensis mtHsp70'"
## [83] "'Tetrahymena borealis mtHsp70'"
## [84] "'Ichthyophthirius multifiliis mtHsp70-2'"
## [85] "'Caenorhabditis elegans mtHsp70'"
## [86] "'Daphnia pulex mt Hsp70'"
## [87] "'Drosophila melanogaster mtHsp70'"
## [88] "'Danio rerio HSPA9B'"
## [89] "'Danio rerio HSPA9B'"
## [90] "'Gallus gallus HSPA9A'"
## [91] "'Homo sapiens HSPA9B'"
## [92] "'Mus musculus HSPA9A'"
## [93] "'Wolbach endosymbiont of Drosophila melanogaster'"
## [94] "'Rhizobium etli DnaK'"
## [95] "'Agrobacterium tumefaciens DnaK'"
## [96] "'Rickettsia prowazekii DnaK'"
## [97] "'Methylobacterium nodulans DnaK'"

```

```

# get species names

```

```

my_group <- my_tree$tip.label

```

```

my_group <- gsub("'", "", my_group)

```

```

my_group_list <- strsplit(my_group, split = " ")

```

```

my_group <- sapply(my_group_list, function(x) paste0(x[1:2], collapse = "_"))

```

```

my_query <- datelife::make_datelife_query(input = my_group)

```

```

## ... Making a DateLife query.

```

```

## ... Phylo-processing 'input'.

```

```

## 'input' is not a phylogeny.
## |
## Warning: Wolbach_endosymbiont are not matched
## |=====
## Working with the following taxa:
## Natronomonas pharaonis | Haloferax volcanii | Methanosarcina barkeri | Methanotherix thermoacetophil
## DateLife query done!
my_query <- datelife::make_datelife_query("mantodea", get_spp_from_taxon = TRUE)

## ... Making a DateLife query.
## ... Phylo-processing 'input'.
## 'input' is not a phylogeny.
## |
## |
## |
## |

str(my_query)

## List of 3
## $ cleaned_names: Named chr [1:362] "Stenophylla gallardi" "Stenophylla lobivertex" "Mantoida schrad
## ..- attr(*, "names")= chr [1:362] "Mantodea1" "Mantodea2" "Mantodea3" "Mantodea4" ...
## $ ott_ids      : Named int [1:362] 4176725 711174 1046541 235229 892960 711168 3484353 1012466 4238
## ..- attr(*, "names")= chr [1:362] "Stenophylla gallardi" "Stenophylla lobivertex" "Mantoida schrad
## $ phy          : logi NA
## - attr(*, "class")= chr "datelifeQuery"
# check how many taxon names you got:
length(my_query$cleaned_names)

## [1] 362

```

DateLife to-go exercise

Instructions:

- 1) Search DateLife's chronogram database to get *ages* for your taxa. To do this, run the function `datelife::get_datelife_result` on your processed list of taxon names (that is, your `my_query` object), assign the output to a named object. I named mine `ages`.
- 2) Summarize the ages as individual study chronograms. Most of the time, published chronograms have other tips that are not part of your taxa of interest. To get just the parts of study chronograms that you are interested in, use the function `datelife::summarize_datelife_result` on the output of the previous step. Make sure to use the argument `summary_format = "phylo_all"`. Assign the output to a different object. I named mine `my_chronograms`.
- 3) install the package `datelifeplot` using the function `devtools::install_github("phylotastic/datelifeplot")`.
- 4) plot your study chronograms with `datelifeplot::plot_phylo_all`. Use option `write = "pdf"` to save your study chronograms as a pdf. They will be saved in a folder named "phylo_all".
- 5) Choose a tree topology to date. You can choose any of the chronograms in `my_chronograms`. You can use the chronogram with the most number of tips. To get that one use the function `datelife::get_biggest_multiphylo`. Alternatively (the recommended option) is to get a synthetic topology from Open Tree for your taxa. To get this one, use function `datelife::get_otol_synthetic_tree` using `my_query` as input.
- 6) Add the topology of your choosing to your `my_query` object. Do this with the following command:
`my_query$phy <- my_topology,`
- 7) Use the age data to date the tree topology that you chose. Use the function `datelife::datelife_use_datelifequery`

My solution:

```
# search DateLife's chronogram database for your taxa
ages <- datelife::get_datelife_result(input = my_query)

# check the structure of the output of datelife::get_datelife_result
names(ages)
str(ages)

# transform matrices back to phylogenetic trees
my_chronograms <- datelife::summarize_datelife_result(datelife_result = ages,
                                                    datelife_query = my_query,
                                                    summary_format = "phylo_all")

# get citations of studies that published your chronograms
names(my_chronograms)

# instal datelifeplot
devtools::install_github("phylotastic/datelifeplot")

# look at the first chronogram:
my_chronograms[[1]]

# check how many taxa we got ages for on the first chronogram:
ape::Ntip(my_chronograms[[1]])
# we got ages for for 91 species
length(my_query$cleaned_names) - ape::Ntip(my_chronograms[[1]])
# we do not have age data for 271 species

# plot your study chronograms and save them as pdf
datelifeplot::plot_phylo_all(chronograms = my_chronograms, write = "pdf")

# Choose a tree topology to date
biggest <- datelife::get_biggest_multiphylo(my_chronograms)
opentree <- datelife::get_otol_synthetic_tree(input = my_query)

# Add it to your my_query$phy
my_query$phy <- opentree

# Date the tree topology, you can use either of the following
dated_opentree <- datelife::datelife_use_datelifequery(datelife_query = my_query)
dated_opentree <- datelife::datelife_use(input = opentree)
```