

# jw\_tree\_opentree

James Waterford

2023-04-20

## Open Tree of Life

There are over 1 million named species of insects.

There are only 5 thousand named species of mammals.

There are about 200 thousand named species of flowering plants.

**APE** stands for ‘Analyses of Phylogenetics and Evolution’

This library helps us read `.tre` files straight from Open Tree of Life.

```
library(ape)
library(rotl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:ape':
##
##      where
```

```
## The following objects are masked from 'package:stats':
##
##      filter, lag
```

```
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
```

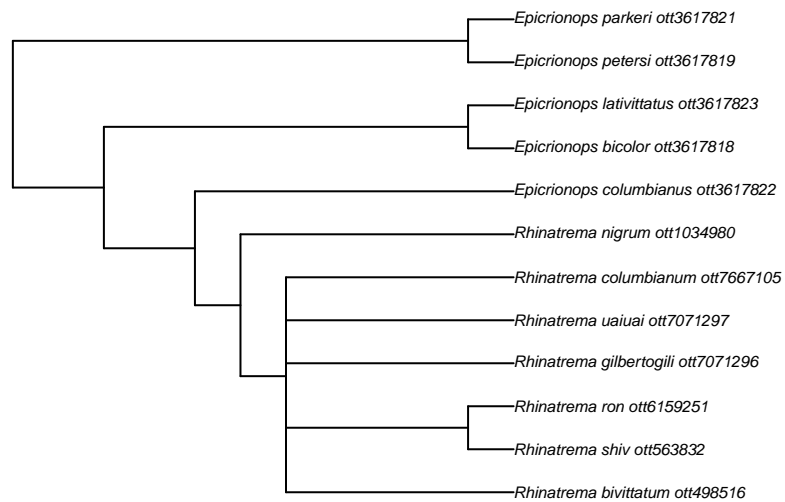
```
library(stringr)
```

Tree was downloaded from:

<https://tree.opentreeoflife.org/opentree/argus/opentree13.4@mrcaott498516ott3617819/Rhinatrema--Epicrionops-parkeri>

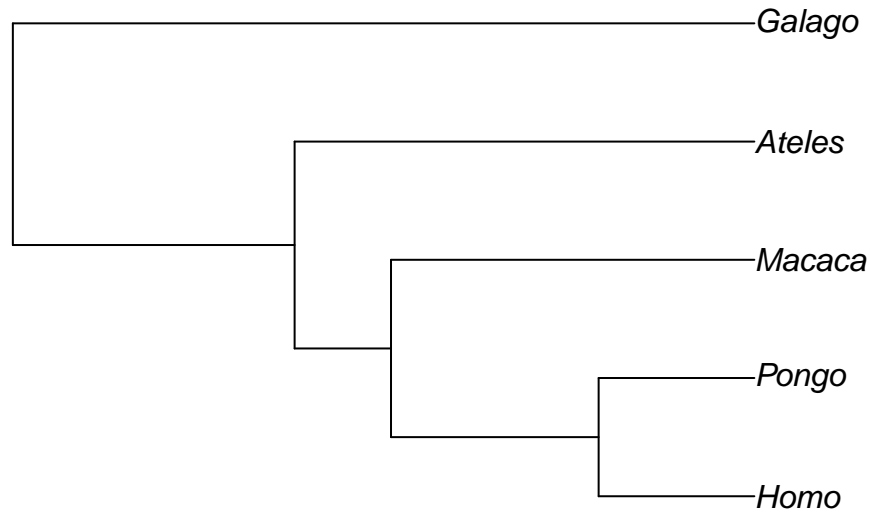
```
tree <- read.tree(file = "../197-raw_storage/subtree-node-mrcaott498516ott3617819-Rhinatrema--Epicrionops-parkeri.tre")
plot(tree, cex = 0.5)
mtext(text = "Walter Jetz 2018 OpenTree tree")
```

## Walter Jetz 2018 OpenTree tree



We can also use a URL to directly call a tree with `ape::`:

```
small_tr <- read.tree(file = "http://ape-package.ird.fr/APER/APER2/primfive.tre")  
plot(small_tr)
```



```
## I need to move these files ##
```

```
a <- "../197-raw_storage/doi_10.5061_dryad.vx0k6djtn_v6/Gleditsch_et_al_Caribbean_Bank_Data.csv"
b <- "../197-raw_storage/doi_10.5061_dryad.vx0k6djtn_v6/Gleditsch_et_al_Caribbean_Herp_SR.csv"
```

```
a_table <- read.csv(a)
b_table <- read.csv(b)
```

```
b_table$Bank <- str_to_title(b_table$Bank)
carribean_set <- right_join(a_table, b_table, by = "Bank")
```

```
## Warning in right_join(a_table, b_table, by = "Bank"): Each row in 'x' is expected to match at most 1
## i Row 1 of 'x' matches multiple rows.
## i If multiple matches are expected, set 'multiple = "all"' to silence this
## warning.
```

```
str(carribean_set)
```

```
## 'data.frame': 1152 obs. of 14 variables:
## $ Bank : chr "Anguilla" "Anguilla" "Anguilla" "Anguilla" ...
## $ Number_of_Islands : int 30 30 30 30 30 30 30 30 30 30 ...
## $ Geographic_Area : num 201 201 201 201 201 ...
## $ Geographic_Isolation : num 750 750 750 750 750 ...
## $ Island_Spread : num 0.848 0.848 0.848 0.848 0.848 ...
## $ Economic_Isolation : int 417 417 417 417 417 417 417 417 417 417 ...
```

```
## $ Average_Population      : num  79505 79505 79505 79505 79505 ...
## $ Topographic_Complexity: num   59.5 59.5 59.5 59.5 59.5 ...
## $ Anthropogenic_Habitat  : num   0.242 0.242 0.242 0.242 0.242 ...
## $ Green_Habitat          : num   0.647 0.647 0.647 0.647 0.647 ...
## $ Clade                   : chr    "All Herpetofauna" "Reptilia" "Squamata" "Iguania" ...
## $ Introduced_SR           : int    18 12 10 4 2 3 2 1 1 2 ...
## $ Native_SR               : int    14 14 14 4 2 2 4 2 1 0 ...
## $ Total_SR                : int    32 26 24 8 4 5 6 3 2 2 ...
```

## The R Structure of a Phylogenetic Tree

### Generating Matching Species Names

We will use `tnrs_match_names()` to find the the proper `Ott_Id` for a given species.

`Ott_id` is the Open Tree of Life Unique Identifier.

```
species <- caribbean_set$Clade[3:16]
species
```

```
## [1] "Squamata"      "Iguania"       "Serpentes"
## [4] "Anolis"        "Gekkota"       "Gymnophthalmoidea"
## [7] "Typhlopidae"   "Testudines"    "Iguanidae"
## [10] "Cryptodira"    "Amphibia"      "Neobatrachia"
## [13] "Eleutherodactylus" "Hylidae"
```

```
caribbean_taxa <-tnrs_match_names(species)
caribbean_taxa
```

```
##           search_string                unique_name approximate_match ott_id
## 1           squamata Squamata (order in Deuterostomia)          FALSE  35888
## 2             iguania                      Iguania          FALSE 608979
## 3           serpentes                      Serpentes          FALSE 186816
## 4             anolis                      Anolis          FALSE 705358
## 5             gekkota                      Gekkota          FALSE 190153
## 6 gymnophthalmoidea          Gymnophthalmidae          TRUE  58583
## 7           typhlopidae          Typhlopidae          FALSE 100036
## 8           testudines          Testudines          FALSE 639666
## 9             iguanidae          Iguanidae          FALSE 608975
## 10          cryptodira          Cryptodira          FALSE  66466
## 11           amphibia          Amphibia          FALSE 544595
## 12          neobatrachia          Neobatrachia          FALSE 535804
## 13 eleutherodactylus          Eleutherodactylus          FALSE 889366
## 14             hylidae                      Hylidae          FALSE 535782
##           is_synonym      flags number_matches
## 1           FALSE                1
## 2           FALSE                1
## 3           FALSE sibling_higher          1
## 4           FALSE                1
## 5           FALSE                1
## 6           FALSE                2
## 7           FALSE                1
```

```
## 8      FALSE      1
## 9      FALSE      1
## 10     FALSE      1
## 11     FALSE      5
## 12     FALSE      1
## 13     FALSE      1
## 14     FALSE      1
```

```
tr <- tol_induced_subtree(ott_ids = carribean_taxa$ott_id)
```

```
## Progress [-----] 0/51 ( 0) ?sProgress [=====]
```

```
## Warning in collapse_singles(tr, show_progress): Dropping singleton nodes
## with labels: Amniota ott229560, Sauropsida ott639642, Sauria ott329823,
## mrcaott246ott4128455, mrcaott246ott4127082, mrcaott246ott4129629,
## mrcaott246ott4142716, mrcaott246ott4126667, mrcaott246ott2982, Testudines
## ott639666, mrcaott1662ott4947157, Lepidosauria ott35881, Squamata (order
## in Deuterostomia) ott35888, Unidentata ott4945815, mrcaott1662ott4126044,
## Serpentes ott186816, mrcaott1662ott20148, mrcaott20148ott56992,
## mrcaott56992ott57008, mrcaott2417ott97368, mrcaott2417ott4124528, Iguania
## ott608979, mrcaott3089ott4125746, mrcaott3089ott4125739, mrcaott3089ott9418,
## mrcaott3089ott53412, mrcaott3089ott6523, mrcaott3089ott15148,
## mrcaott3089ott7982, mrcaott3089ott116302, mrcaott3089ott32977,
## mrcaott3089ott10227, mrcaott4987ott10058, mrcaott4987ott41163, Amphibia
## ott544595, Batrachia ott471197, Anura ott991547, mrcaott114ott3129,
## mrcaott114ott37876, mrcaott114ott18818, Neobatrachia ott535804,
## mrcaott114ott309463, mrcaott114ott177, mrcaott177ott7464, mrcaott177ott29310,
## mrcaott177ott1321, mrcaott2199ott411156, mrcaott2199ott107387,
## mrcaott2199ott22795, Eleutherodactylidae ott63340, Eleutherodactylinae ott478714
```

```
plot(tr)
```

