

Tree from Open Tree

Luna L Sanchez Reyes

2023-04-20

Insects - 1 million named species Mammals - 5k named species Flowering plants - 200k named species

Read a newick tree from the Open Tree of Life

The following function can only read trees in newick format. Make sure your file is a newick tree. If you got your tree from the Phylsystem repo, your tree is going to be in json format and will not be read by the function.

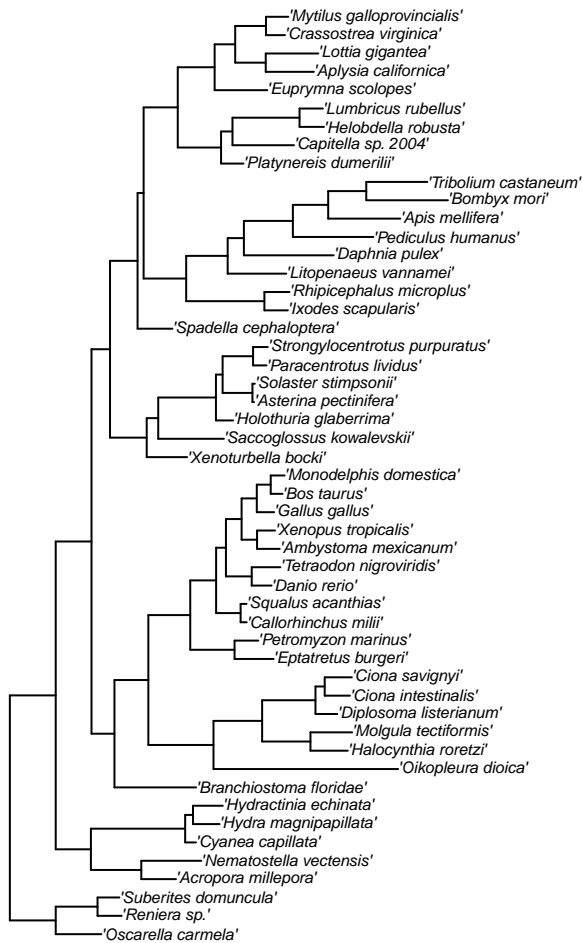
Newick Tree downloaded from <https://tree.opentreeoflife.org/opentree/opentree13.4@ott118027/Gymnophiona>

```
tree <- read.tree(file = "../data-raw/delsuc_2008.tre")
```

Now plot the tree:

```
plot(tree, cex = 0.5)  
mtext(text = "Delsuc 2008 OpenTree tree")
```

Delsuc 2008 OpenTree tree



We can also use the url of a tree to read a tree into R. We will provide the URL in the argument `file` =. In this case, we do not have to keep a copy of the file, we can just download it directly. The url is "<http://ape-package.ird.fr/APER/APER2/primfive.tre>":

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

```
##
## Phylogenetic tree with 5 tips and 4 internal nodes.
##
## Tip labels:
##   Homo, Pongo, Macaca, Ateles, Galago
##
## Rooted; includes branch lengths.
```

The R structure of a phylogenetic tree

Getting a tree for your species from the Open Tree of Life

The Open Tree of Life has tools that allow to match the names of a group or species to their unified taxonomy so that we can look for information on their databases.

If we are doing this in R, we are using the `tnrs_match_names()` from the package `rotl`; `tnrs` stands for taxonomic name resolution service:

```
tnrs <- tnrs_match_names(names = "amphibians")
class(tnrs)
```

```
## [1] "match_names" "data.frame"
```

```
tnrs
```

```
##   search_string unique_name approximate_match ott_id is_synonym flags
## 1   amphibians   Amphibia                TRUE 544595      FALSE
##   number_matches
## 1                6
```

What is the main difference between the information that we have in the column `search_string` and `unique_name`: One is the search string and the other is the scientific taxonomic name from the Open Tree of Life unified Taxonomy (OTT).

We can use the unique number identifier from OTT to get information from that taxon.

To get a tree we can use the function `tol_subtree()`

```
tol_subtree(ott_id = tnrs$ott_id)
```

```
## Progress [-----] 0/189 ( 0) ?sProgress [=====]

## Warning in collapse_singles(tr, show_progress): Dropping singleton nodes
## with labels: Rana clamitans ott515378, Rana sphenoccephala ott61437, Rana
## aurora ott771201, Glandirana ott407917, Pterorana ott3618410, Pelophylax
## cf ott7070897, Hemimantis ott4133632, Microdiscopus ott4133623, Montorana
## ott4133633, Chiromantis vittatus ott389176, Mercurana ott4133643, Beddomixalus
## ott4133641, Buergeriinae ott223222, Maitisomantis ott678997, Boehmantis
## ott484429, laevigata group ott189959, bernhardi group ott189958, Wakea
## ott484449, Tsingymantis ott976861, Boophinae ott764200, Laliostoma ott1054409,
## Ombrana ott4133654, Chrysopaa ott4133659, Alcalinae ott5926144, Liurananinae
## ott5926192, Natalobatrachus ott532110, Nothophryne ott3618085, Cacosternum
## nanum ott676305, Microbatrachella ott751405, Poyntonia ott475120, Anhydrophryne
## ott113819, Ericabatrachus ott3618087, Phrynobatrachidae ott504589, Micrixalidae
## ott1081209, Odontobatrachidae ott5536254, Lanzarana ott3618164, Trichobatrachus
## ott780979, Pararthroleptis ott4133674, Hyperolius fusciventris ott85268,
## Chlorolius ott3619081, Tachycnemis (genus in Deuterostomia) ott750024, Morerella
## ott85260, Kassina ott3619075, Opisthophryne ott645881, Chrysobatrachus
## ott3619077, Callixalus ott3619073, Arlequinus ott3619083, Semnodactylus
## ott62328, Tornierella ott4133669, Breviceps adspersus ott3618690, Spelaeophryne
## ott660153, Probreviceps macrodactylus ott111206, Balebreviceps ott348136,
## Hemisotidae ott165721, Oninia ott789802, Genyophryne ott21522, Siamophryne
## ott7070469, Melanobatrachus ott701612, Kalophrynidae ott977735, Phrynellia (genus
## in Opisthokonta) ott401609, Mysticellus ott7070618, Dermatohyla ott186184,
## Gastrophryne olivacea ott565409, Arcovomer ott844415, Adelastinae ott5800508,
## Adelastes ott3618939, Dasypops ott513414, Relictivomer ott190096, Anilany
## ott5926118, Madecassophryne ott3618941, Parhoplophryne ott3618937, Otophryne
## ott404351, Phrynomerinae ott630304, Caluella ott7666463, Corythomantis
## ott442032, Argenteohyla ott578372, Argenteohyla siemersi ott100569, Nyctimantis
## ott1087156, Itapotihyla ott257368, Hyla annectans ott655531, Anotheca ott59159,
## Diagonalia ott3620134, Acris crepitans ott59141, Quilticohyla ott7070260, Rheohyla
## ott7070262, Nesorohyla ott7070190, Pachymedusa ott254792, Dryaderces ott7666055,
## Didynamipus ott152264, Barbarophryne ott5800473, Churamiti ott104959, Bufotes
## pewzowi ott1072351, Schismaderma ott506368, Sabahphryne ott4133471, Strauchbufo
## ott6158681, Epidaleia ott334615, Anaxyrus americanus ott889326, Laurentophryne
## ott3619734, Parapelophryne ott3619737, Pseudobufo ott3619739, Bufoides
```

```

## ott3619731, Metaphryniscus ott3619751, Blythophryne ott5926045, Ghatophryne
## ott7069948, Rentapia ott7069968, Silverstoneia nubicola ott638061, Hyloxalinae
## ott1096759, Lithodytes ott315881, Scythrophrys ott462991, Rupirana ott3620326,
## Phrynocerus ott3619498, Physalalemus ott6158772, Niedenia ott4133301,
## Allophrynidae ott57740, Macrogenioglottus ott726713, Insuetophrynus ott44381,
## Telmatobiinae ott777187, Hylorina ott440882, Chaltenobatrachus ott6158620,
## Limnomedusa ott914517, Chacophrys ott431943, Caudiverbera ott1068325,
## Atopophrynus ott3620196, Geobatrachus ott3620197, Ceuthomantinae ott277741,
## Megistolotis ott276283, Adelotus ott276281, Assa (genus in Opisthokonta)
## ott906692, Paracrinia ott989413, Metacrinia ott412471, Spicospina ott1039925,
## Rheobatrachidae ott918183, Hadromophryne ott971912, Atympophrys ott542885,
## Vibrissaphora ott535112, Pelodytidae ott509554, Xenopodinae ott940173,
## Pseudhymenochirus ott140873, Rhinophrynidae ott459016, Rhinophrynus ott459015,
## Alytinae ott5334814, Discoglossus galganoi ott461369, Latonia ott4948216,
## Leiopelmatidae ott611960, Ascaphidae ott1013114, Pelodryadidae ott3620482,
## Ranoidea (genus in family Pelodryadidae) ott7666856, Iranodon ott7071233,
## Ranodon ott834698, Satobius ott5800418, Pachyhynobius ott1021854, Ambystomatidae
## ott984723, Dicamptodontidae ott60819, Laotriton ott4948201, Triturus marmoratus
## ott1041767, Triturus carnifex ott1041783, Ommatotriton ophryticus ott645229,
## Lissotriton helveticus ott9366, Lissotriton boscai ott830424, Tylototriton
## verrucosus ott932561, Tylototriton wenxianensis ott981376, Pleurodeles waltl
## ott566038, Lyciasalamandra fazilae ott1024882, Salamandrininae ott4948210,
## Pingia ott4132654, Isthmura bellii ott46162, Isthmura sierraoccidentalis
## ott98841, Parvimolge ott46159, Bradytriton ott798666, Nyctanolis ott224306,
## Haideotriton ott133635, Eurycea multiplicata ott839995, Stereochilus (genus
## in Opisthokonta) ott798664, Hemidactylium ott798658, Phaeognathus ott964128,
## Karsenia ott893551, Ensatina ott64118, Hydromantes imperialis ott675243,
## Atylodes (genus in Deuterostomia) ott693464, Hydromantoides ott4132648,
## Amphiumidae ott566022, Rhyacotritonidae ott459010, Sirenoidea ott336754,
## Chikilidae ott4948197, Praslinia ott80530, Sylvaecilia ott3617918,
## Atretochoana ott3617929, Potamotyphlus ott7667119, Bdellophis ott4132629

##
## Phylogenetic tree with 10020 tips and 4669 internal nodes.
##
## Tip labels:
## Odorrana_geminata_ott114, Odorrana_chapaensis_ott214633, Odorrana_grahami_ott43280, Odorrana_marga
## Node labels:
## Amphibia ott544595, Batrachia ott471197, Anura ott991547, , , , ...
##
## Unrooted; no branch lengths.

```

Exercise

Get a subtree from a group using the function `tol_subtree()`; extract the OTT id using the function `tnrs_match_names()`.

```

tnrs_dolphins <- tnrs_match_names(names = "dolphins")
tnrs_dolphins

```

```

## search_string unique_name approximate_match ott_id is_synonym flags
## 1 dolphins Delphinus TRUE 698409 FALSE sibling_higher
## number_matches
## 1 1

```

The function above ran the TNRS service to match the common name that I provided to the Open Tree

of Life Taxonomy (OTT) and get a unique numeric identifier called the OTT id. For dolphins this is the number 698409. The next step is to use the OTT id in the function that gets a subtree from Open Tree:

```
tol_subtree(ott_id = tnrs_dolphins$ott_id) -> dolphins_subtree
```

```
## Progress [-----] 0/1 ( 0) ?sProgress [=====]
```

```
## Warning in collapse_singles(tr, show_progress): Dropping singleton nodes with
## labels: Delphinus delphis ott698438
```

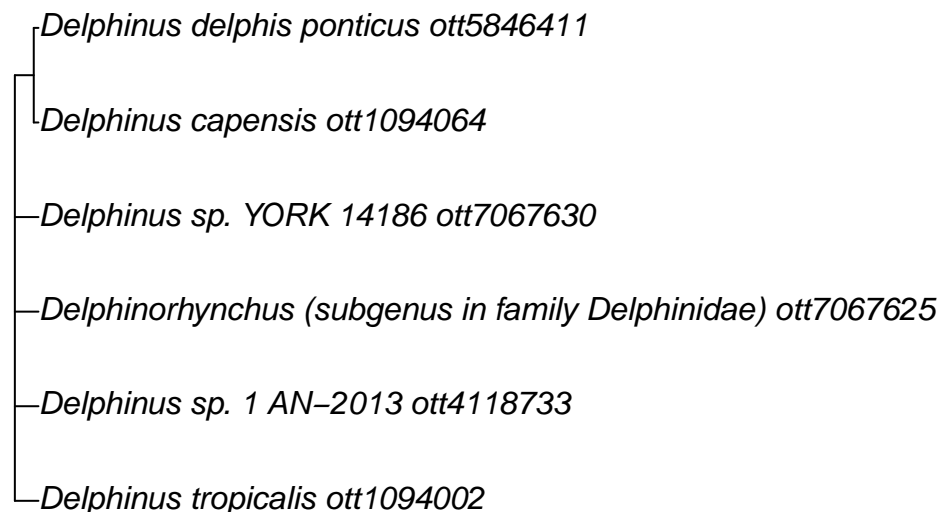
```
str(dolphins_subtree)
```

```
## List of 4
## $ edge      : int [1:7, 1:2] 7 8 7 7 7 8 1 2 4 ...
## $ tip.label : chr [1:6] "Delphinus_tropicalis_ott1094002" "Delphinus_capensis_ott1094064" "Delphinus
## $ Nnode      : num 2
## $ node.label: chr [1:2] "Delphinus ott698409" ""
## - attr(*, "class")= chr "phylo"
```

```
class(dolphins_subtree)
```

```
## [1] "phylo"
```

```
plot.phylo(x = dolphins_subtree)
```



```
Ntip(dolphins_subtree)
```

```
## [1] 6
```

Getting a tree for a set of taxa

Sometime we have very large groups and we just want to get a small tree representing some lineages inside the group.

For example, we just want a tree of all orders in amphibians. We know the common names of these which are frogs, salamanders and cecilians.

We want to find this lineages in the Open Tree of Life; the steps to do this are:

1. Get the scientific names of our taxa

```
c("anura", "urodela", "gymnophiona") -> amphibian_orders
```

2. Match the scientific names to the OTT to get a unique numeric identifier

```
tnrs_match_names(names = amphibian_orders) -> amph_order_tnrs
amph_order_tnrs
```

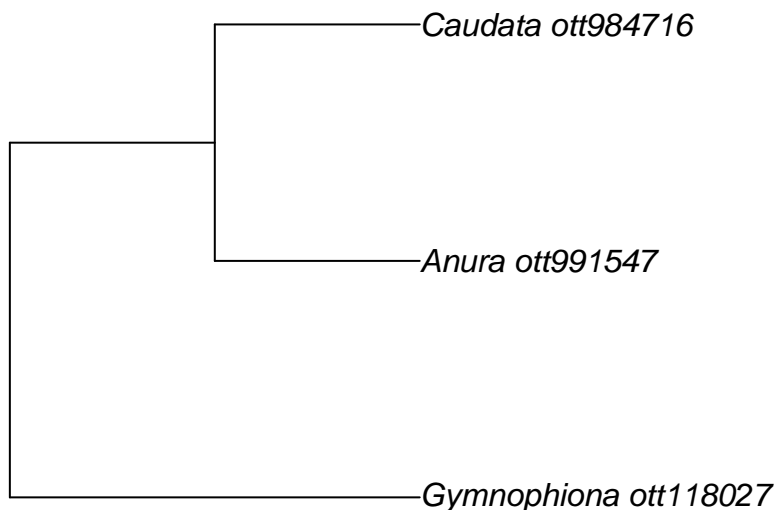
```
## search_string unique_name approximate_match ott_id is_synonym flags
## 1 anura Anura FALSE 991547 FALSE
## 2 urodela Caudata FALSE 984716 TRUE
## 3 gymnophiona Gymnophiona FALSE 118027 FALSE sibling_higher
## number_matches
## 1 6
## 2 2
## 3 1
```

We have two columns at the beginning, search string contains the names that we provided, and unique_name contains the most updated or official or most commonly used taxonomic names for the search taxon that we provided.

3. Next step is to use the OTT ids that we got to extract the tree containing those lineages. If we have only one name or lineage, we can use `tol_subtree()`. If we have more than one names, we will use `tol_induced_subtree()`.

```
tol_induced_subtree(ott_ids = amph_order_tnrs$ott_id) -> amp_order_tree
amp_order_tree
```

```
##
## Phylogenetic tree with 3 tips and 2 internal nodes.
##
## Tip labels:
## Anura_ott991547, Caudata_ott984716, Gymnophiona_ott118027
## Node labels:
## Amphibia_ott544595, Batrachia_ott471197
##
## Rooted; no branch lengths.
plot.phylo(x = amp_order_tree)
```



Getting a tree of all families within amphibians

Install the `datelife` R package from GitHub, with a function from the package `remotes`:

```
library(remotes) # or install.packages("remotes")
install_github(repo = "phylotastic/datelife")
```

Skipping install of 'datelife' from a github remote, the SHA1 (5dde3784) has not changed since last
Use `force = TRUE` to force installation

Now we can use the `datelife` package, but we have to load it first

```
library(datelife)
```

We are going to use a function that extracts all the names from a taxon that belong to a given named rank.
The function is called `get_ott_children()`:

```
##?get_ott_children
get_ott_children(ott_ids = trns$ott_id, ott_rank = "family") -> amphibian_families
```

```
## |
## |
## |
## |
## |
## |
```

```
amphibian_families
```

```
## $Amphibia
##           ott_id  rank
## Caeciliidae    118029 family
## Typhlonectidae  639647 family
## Ichthyophiidae  639653 family
## Dermophiidae   654645 family
## Rhinatrematidae 128153 family
## Indotyphlidae   114139 family
## Siphonopidae    114359 family
## Scolecomorphidae 861429 family
## Herpelidae      379929 family
## Chikilidae      4948197 family
## Leiopelmatidae  611960 family
## Ascaphidae      1013114 family
## Bombinatoridae  1013112 family
## Alytidae        991545 family
## Plethodontidae  515329 family
## Proteidae       630300 family
## Dicamptodontidae 60819 family
## Rhyacotritonidae 459010 family
## Salamandridae   566011 family
## Amphiumidae     566022 family
## Ambystomatidae  984723 family
## Hynobiidae       336749 family
## Cryptobranchidae 336750 family
## Sirenidae        515352 family
## Pelodytidae     509554 family
## Megophryidae    828160 family
## Pelobatidae     485821 family
## Rhinophrynidae  459016 family
## Pipidae         465087 family
```

```
## Hemiphractidae      245330 family
## Leptodactylidae    414584 family
## Heleophrynidae     127159 family
## Alsodidae          533096 family
## Cycloramphidae     533097 family
## Hylidae            535782 family
## Strabomantidae     199520 family
## Brachycephalidae   719424 family
## Rhinodermatidae    4133476 family
## Rheobatrachidae    918183 family
## Hylodidae          1059295 family
## Eleutherodactylidae 63340 family
## Craugastoridae     63341 family
## Dendrobatidae      581837 family
## Bufonidae          889358 family
## Ceratophryidae     1008932 family
## Batrachylidae      1008933 family
## Petropedetidae     337155 family
## Ceratobatrachidae  1081207 family
## Micrixalidae       1081209 family
## Mantellidae        38969 family
## Nyctibatrachidae   1081210 family
## Ptychadenidae      1081211 family
## Ranidae            364560 family
## Ranixalidae        403946 family
## Rhacophoridae      432783 family
## Phrynobatrachidae  504589 family
## Pyxicephalidae     504591 family
## Odontobatrachidae  5536254 family
## Dicroglossidae     1081208 family
## Myobatrachidae     940181 family
## Calyptocephalellidae 346415 family
## Nasikabatrachidae  542087 family
## Sooglossidae       882346 family
## Hyperoliidae       535796 family
## Hemisotidae        165721 family
## Brevicipitidae     790993 family
## Microhylidae       1062202 family
## Arthroleptidae     968547 family
## Alloprynidae       57740 family
## Centrolenidae      467442 family
```

Use the ott ids for all families to get an induced subtree:

```
tol_induced_subtree(ott_ids = amphibian_families$Amphibia$ott_id) -> amp_fam_subtree
```

```
## Progress [-----] 0/15 ( 0) ?sProgress [=====]

## Warning in collapse_singles(tr, show_progress): Dropping singleton nodes with
## labels: mrcaott114ott391676, mrcaott15857ott152667, mrcaott270630ott3618180,
## mrcaott22583ott100573, mrcaott22583ott44382, mrcaott44382ott72638,
## mrcaott44382ott100564, mrcaott65695ott254163, mrcaott65695ott121259,
## mrcaott2199ott411156, mrcaott7464ott21502, mrcaott21502ott918196, Pelobatoidea
## ott485820, mrcaott18818ott47772, Sirenoidea ott336754
```



```
amp_fam_subtree
```

```
##
```

```
## Phylogenetic tree with 60 tips and 59 internal nodes.
```

```
##
```

```
## Tip labels:
```

```
##   Ranidae_ott364560, Rhacophoridae_ott432783, Mantellidae_ott38969, Ranixalidae_ott403946, Nyctibatr
```

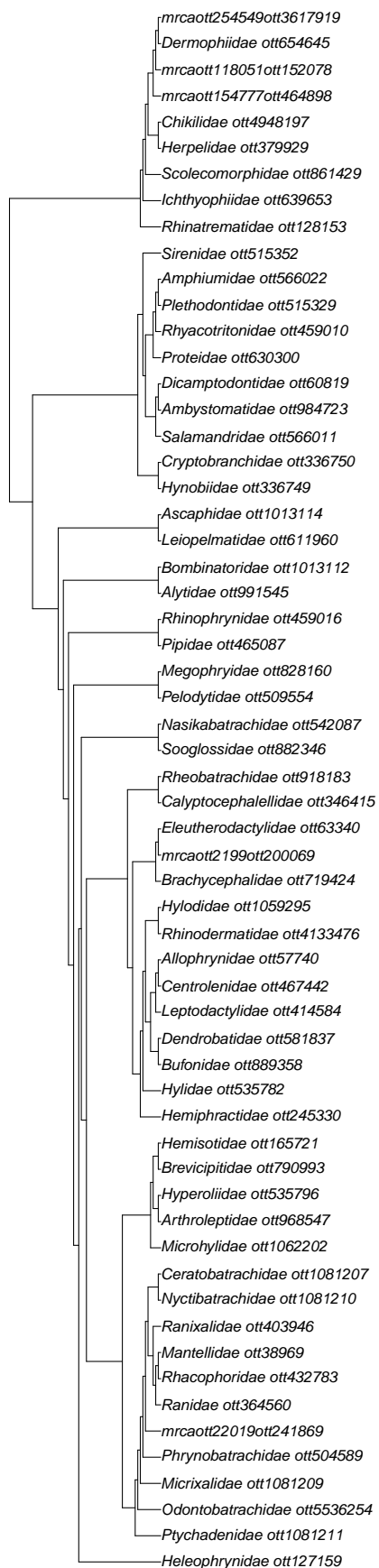
```
## Node labels:
```

```
##   Amphibia ott544595, Batrachia ott471197, Anura ott991547, mrcaott114ott3129, mrcaott114ott37876, m
```

```
##
```

```
## Rooted; no branch lengths.
```

```
plot.phylo(amp_fam_subtree)
```



Exercise:

1. Extract the names of the taxa from your data set for the final project.
2. Match them to the OTT using TNRS with the function `tnrs_match_names()`
3. Use the OTT ids to get an induced subtree with the function `tol_induced_subtree()`.

I'll use the CSV file of the diet of *Canis lupus*, that we got from here <https://datadryad.org/stash/dataset/doi:10.5061%2Fdryad.zgmsbcegg>, as an example

```
read.csv(file = "../data-raw/scat_RRA.csv") -> scat
head(scat)
```

##	SampleNumber	BioYear	pack	hr_type	Prey	mean_read_count
## 1	6	2014	Honker	annual	Odocoileus hemionus	24852.5
## 2	7	2014	Honker	annual	Odocoileus hemionus	112592.0
## 3	8	2014	Snow Pass	annual	Odocoileus hemionus	3585.5
## 4	8	2014	Snow Pass	annual	Lontra canadensis	1475.5
## 5	8	2014	Snow Pass	annual	Phoca vitulina	697.0
## 6	9	2014	Trocadero	annual	Odocoileus hemionus	40730.0
##	total_reads_sample	prop_reads_sample				
## 1	24852.5	1.0000000				
## 2	112592.0	1.0000000				
## 3	5758.0	0.6226989				
## 4	5758.0	0.2562522				
## 5	5758.0	0.1210490				
## 6	40730.0	1.0000000				

We do not really have access to the subspecies names of the wolves, only their packs. So we can start by creating a tree for the preys. First, extract the scientific names:

```
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
```

```
scat$Prey %>%
  unique() %>%
  tnrs_match_names() -> prey_ott_tnrs # match the scientific names to the Open Tree of Life Taxonomy
```

```
## Warning: Pica sp., Mictrotus sp. are not matched
```

```
prey_ott_tnrs %>%
  filter(!is.na(ott_id)) -> prey_ott_tnrs
```

```
tol_induced_subtree(ott_ids = prey_ott_tnrs$ott_id)
```

```
## Error: HTTP failure: 400
```

```
## [/v3/tree_of_life/induced_subtree] Error: node_id 'ott5857335' was not found!list(ott173070 = "pruned")
```