

Lab 1: Introduction to Phylogenies in R

Comparative Biology and Macroevolution

April 5, 2019

Learning Objectives

After this lab, you should be able to:

- Install and load packages in R for phylogenetic analysis
- Read and manipulate trees in R
- Plot trees in R using functions from both **ape** and **ggtree**
- Begin thinking of a clade that interests you for your final project

Installing packages

For this class, we will be primarily using R to manipulate and analyze phylogenetic trees. When you download R from the Comprehensive R Archive Network (CRAN), you get the base R system. The advantage of using R is that there are a lot of packages that other researchers have developed that extends the basic functionality in R. In particular, we will focus on the R packages developed for phylogenetic and comparative approaches.

Today, we will be using two R packages, **ape** and **ggtree** to read in and manipulate trees. **ggtree** requires R version 3.5.3. Therefore, you if you own a mac, you need to have OS X 10.11 (El Capitan) or higher.

```
# To install ape
install.packages("ape")

# To install ggtree
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("ggtree", version = "3.8")

# To install ggimage
install.packages("ggimage")
```

After you install a package, you must load it to access its functions.

```
# To load libraries (packages)
library(ape)
library(ggtree)
library(ggimage)
```

Now we can utilize the functions within both libraries. Let's try reading in a phylogenetic tree using one of the functions in the ape package

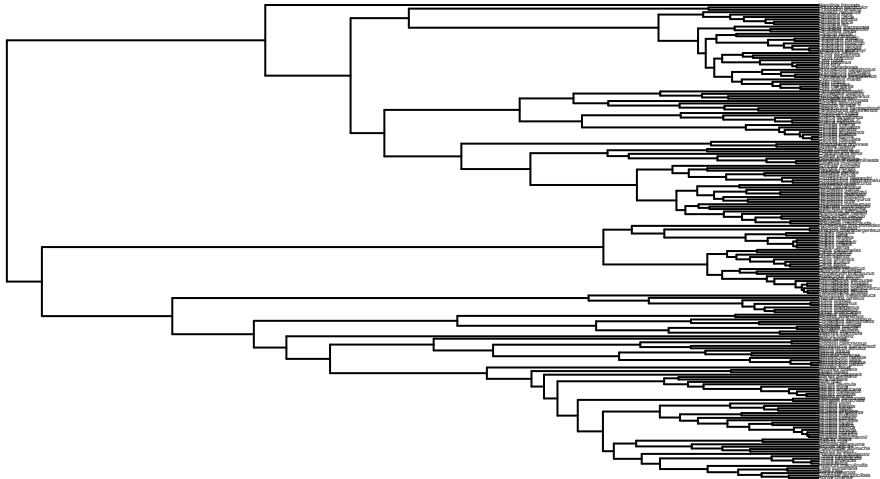
```
load("Carnivoran_Tree")
```

Carnivoran_tree is a list of the class *phylo*. There are at least 3 elements: `edge`, `tip.label`, and `Nnode`.

```
Carnivoran_Tree # Tree has 198 species
```

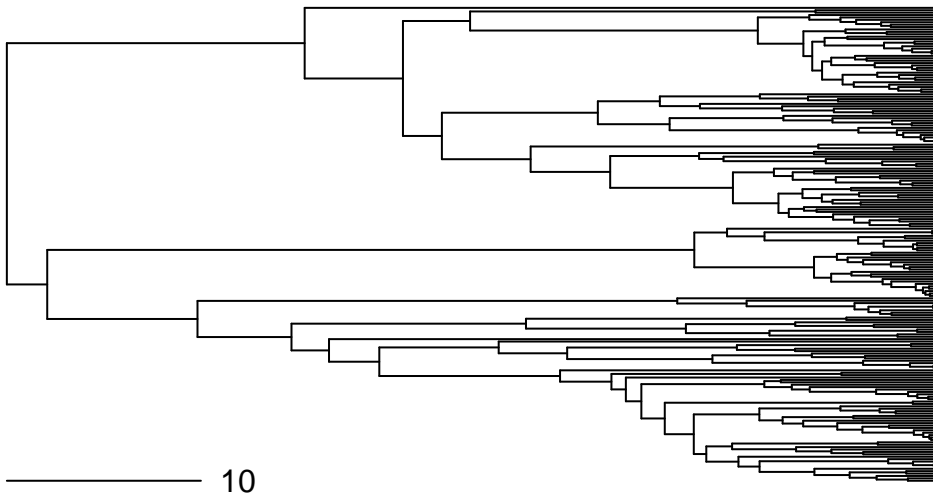
```
##
```

```
## Phylogenetic tree with 198 tips and 197 internal nodes.
```

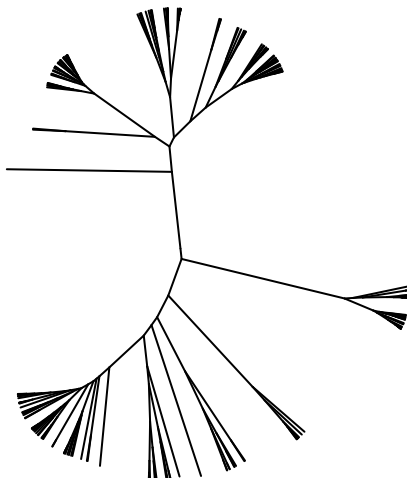
You can also remove the tip labels and add a scale bar to the plot.

```
plot(Carnivoran_Tree, show.tip.label = F)
add.scale.bar()
```



There are different types of phylogenies that you can plot

```
plot(Carnivoran_Tree, show.tip.label = F, type = "unrooted")
```

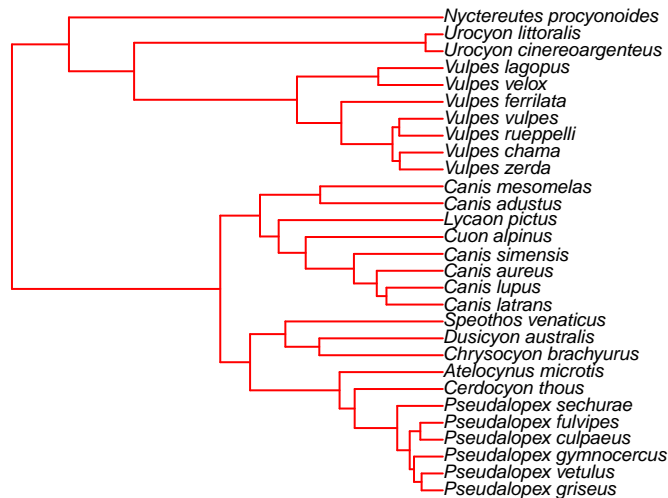


```
plot(Carnivoran_Tree, show.tip.label = F, type = "fan")
```



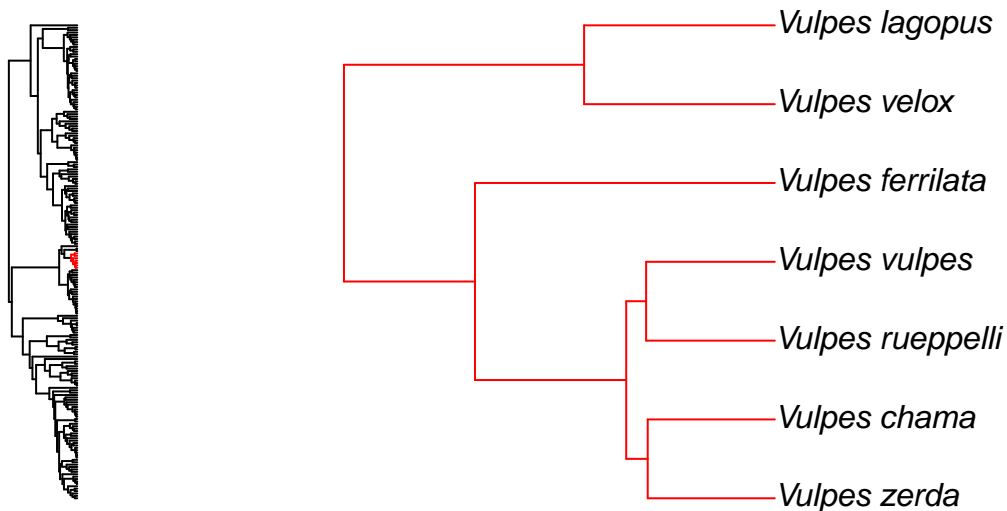
Zoom on a portion of the tree using the `zoom` function in *ape*. We can use the `Family` element of `Carnivoran_Tree` to pull out the family Canidae (dogs).

```
zoom(Carnivoran_Tree, grep("Canidae", Carnivoran_Tree$Family), cex = 0.6)
```



We can then zoom in even more and only look at the genus *Vulpes* by searching using the `tip.label` element of `Carnivoran_Tree`

```
zoom(Carnivoran_Tree, grep("Vulpes", Carnivoran_Tree$tip.label))
```



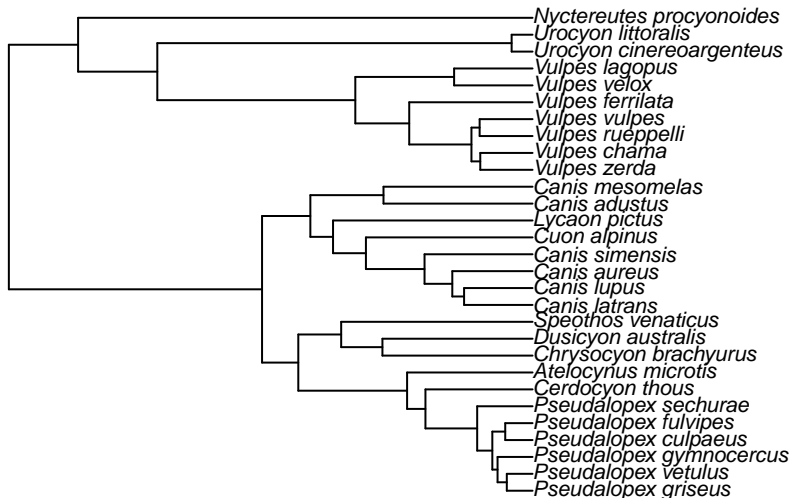
Going back to the Ursidae tree, we can extract the family from the larger Carnivoran phylogeny. First we need to find the node number for the most recent common ancestor of the Ursidae tree.

```
MRCA <- getMRCA(Carnivoran_Tree, c("Nyctereutes procyonoides", "Pseudalopex griseus"))
MRCA
```

```
## [1] 277
```

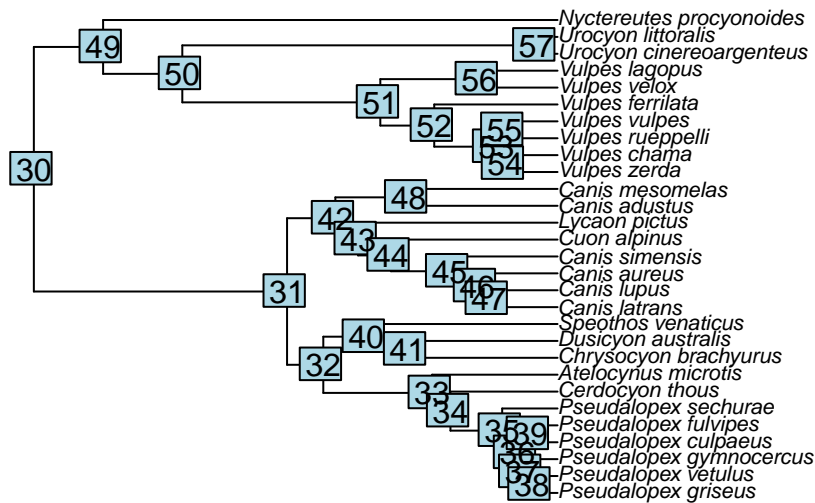
Now we can extract the clade of interest using the node number (277 in this case)

```
Canidae_Tree <- extract.clade(Carnivoran_Tree, node = MRCA)
plot(Canidae_Tree, cex = 0.7)
```

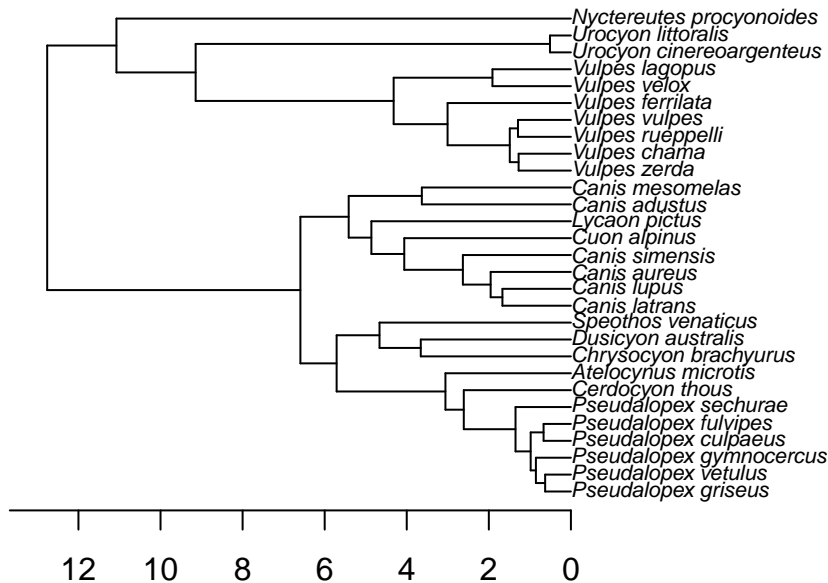


Canidae_Tree is a new *phylo* object

```
plot(Canidae_Tree, cex = 0.7)
nodelabels()
```



```
plot(Canidae_Tree, cex = 0.7)
axisPhylo()
```



Great! Now we can save this tree by writing the extracted tree to a new file.

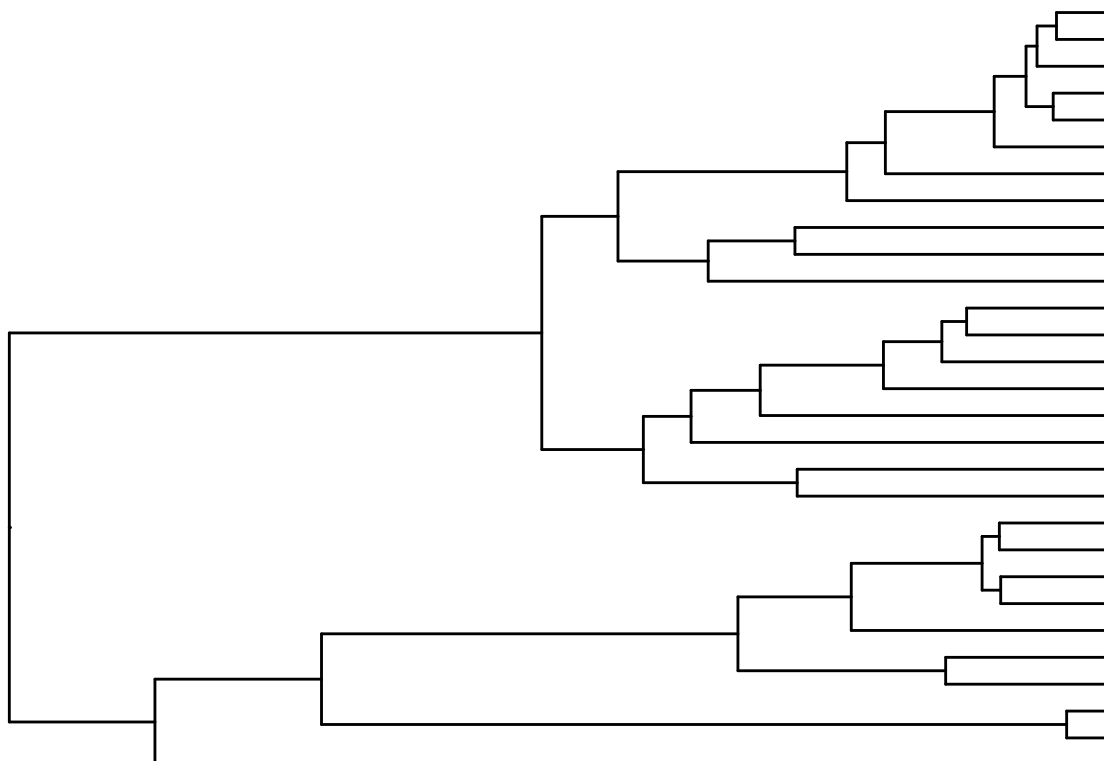
```
write.tree(Canidae_Tree, "Canidae_Tree")
```

Also, write a .csv file with tip.label

```
Canidae_Data <- Canidae_Tree$tip.label
write.csv(Canidae_Data, file = "Canidae_Data.csv")
```

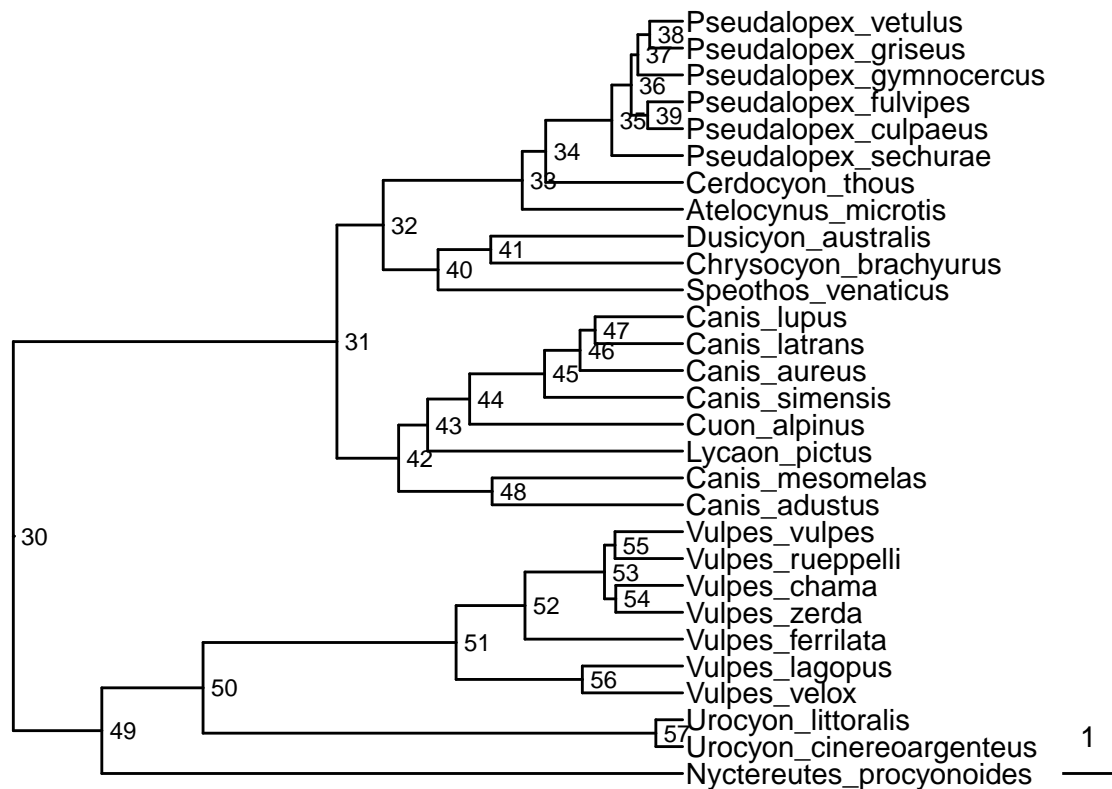
We can also use *ggtree* to plot phylogenetic trees in R.

```
ggtree(Canidae_Tree)
```



You can highlight clades within your tree using *ggtree*

```
CT <- ggtree(Canidae_Tree) + geom_tiplab() +
  geom_treescale(x=20, y=1, offset=1) +
  geom_text2(aes(subset=!isTip, label=node), hjust=-.3, size = 3)
print(CT)
```



```
CladeA <- getMRCA(Canidae_Tree, tip = c("Canis_adustus", "Pseudalopex_vetulus"))
CladeA
```

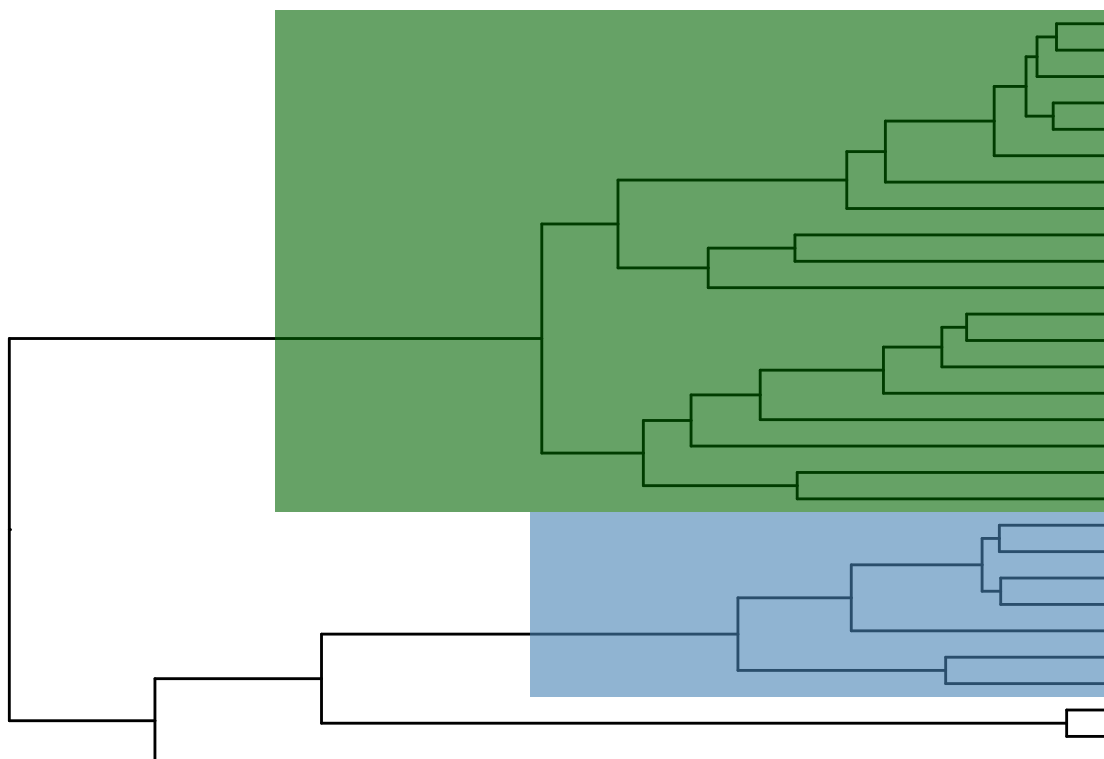
```
## [1] 31
```

```
CladeB <- getMRCA(Canidae_Tree, tip = c("Vulpes_vulpes", "Vulpes_velox"))
CladeB
```

```
## [1] 51
```

We can also see that the node labels from `getMRCA` match the node labels from the plot.

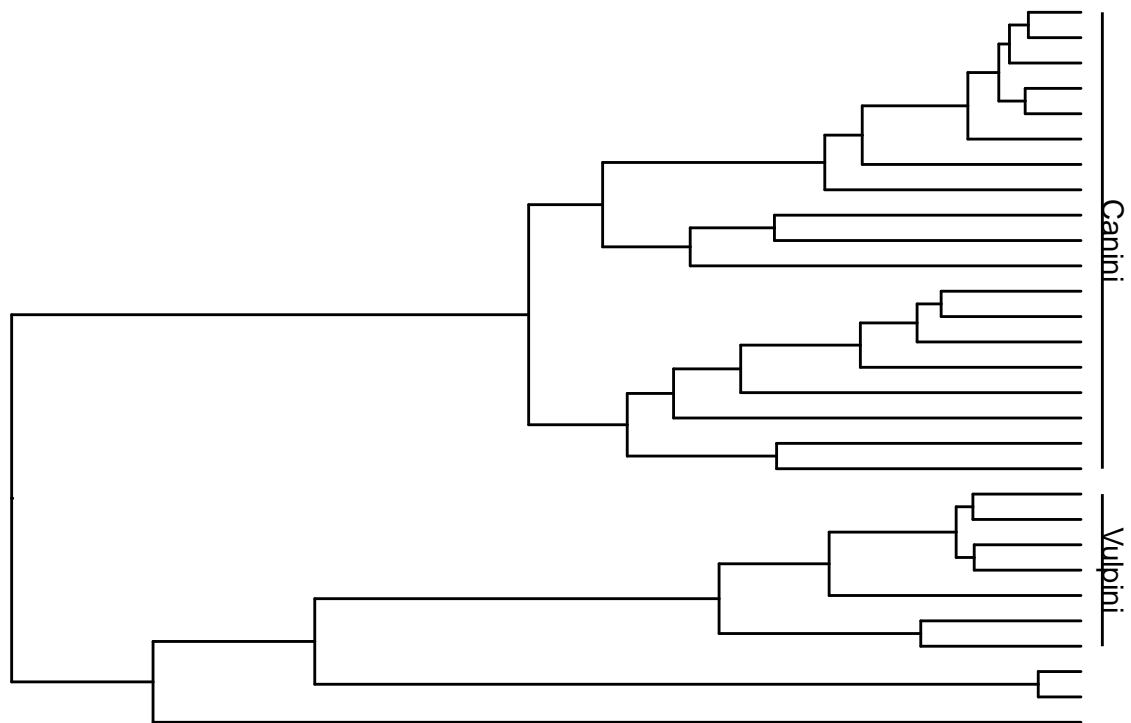
```
CT <- ggtree(Canidae_Tree) +
  geom_highlight(CladeA, fill = "darkgreen", alpha = 0.6) +
  geom_highlight(CladeB, fill = "steelblue", alpha = 0.6)
print(CT)
```

You can also annotate clades of your tree using *ggtree*

```
CT <- ggtree(Canidae_Tree) +  
  geom_cladelabel(node = CladeA, label = "Canini", align=T, angle=270, hjust='center') +  
  geom_cladelabel(node = CladeB, label = "Vulpini", align=T, angle=270, hjust='center') +  
  ggtitle("Canidae Phylogeny")  
print(CT)
```

Canidae Phylogeny

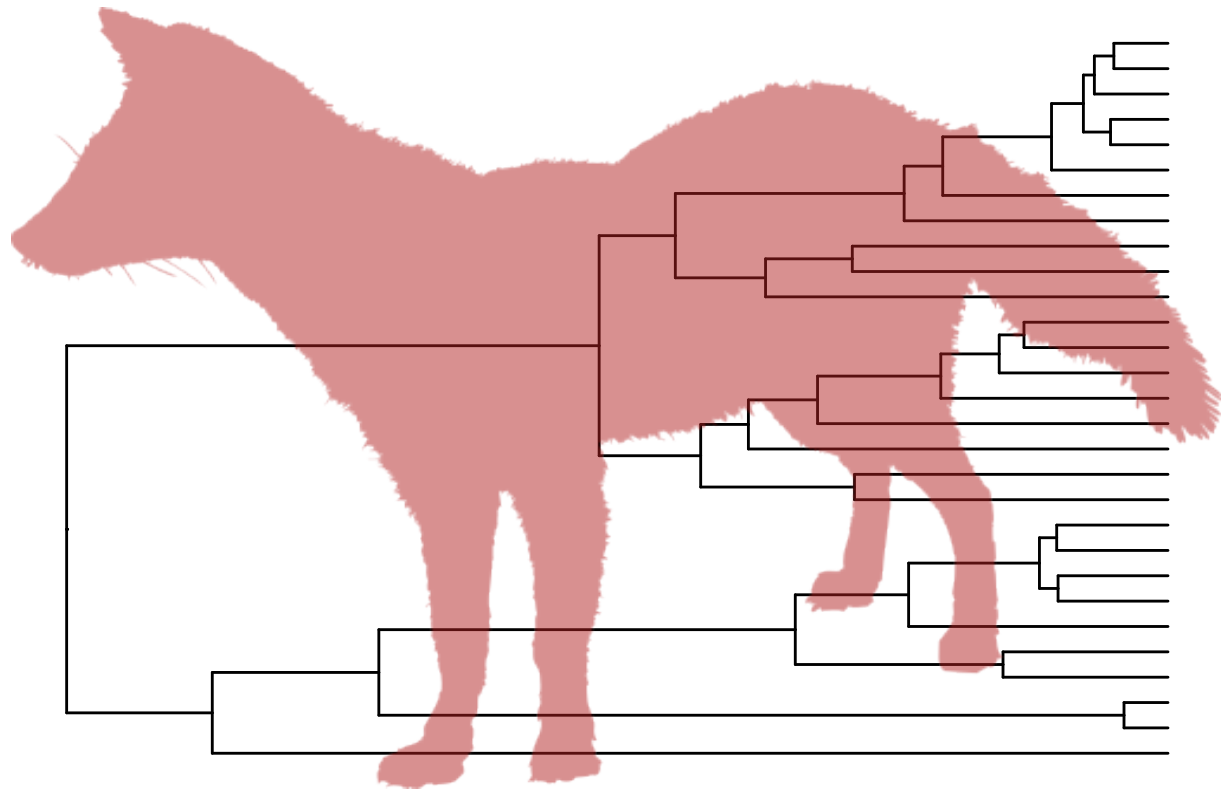


Finally, you can find images to add to your tree on PhyloPic <http://phylopic.org/image/20da6c7c-2584-4cee-921b-ebd09384567b>

Download the phylopic image (Here I saved it as “canidae.png”)

```
CT <- ggtree(Canidae_Tree) +  
  geom_image(image = "canidae.png",  
            size=Inf, alpha=.5,  
            color='firebrick') +  
  ggtitle("Canidae Phylogeny")  
print(CT)
```

Canidae Phylogeny



Homework (Due before lab next week)

Your assignment this week is to produce the 4 figures: You do not need to write a lab report. Please submit a single document containing the 4 plots. All plots should have appropriately sized text (if required) which should fit in the figure.

1. Find a clade of at least 100 species that interests you. Search by tip.label in phylo objects `bird.tre` (Cooney et al. 2017), `fish.tre` (Rabosky et al. 2019), `mammal.tre` (Fritz et al. 2009), `reptile.tre` (Tonini et al. 2016), `ant.tre` (Nelsen et al. 2018) and `plant.tre` (Qian and Jin 2016). Zoom on your clade of interest and plot the result (i.e., subtree next to mega-phylogeny).

Hint, to read in these trees, use the function `read.tree` in `ape`

Example

```
mammal_tree <- read.tree("mammal.tre")
class(mammal_tree)
```

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

```
mammal_tree
```

```
##
```

```
## Phylogenetic tree with 5020 tips and 5019 internal nodes.
```

```
##
```

```
## Tip labels:
```

```
## Rattus_tiomanicus, Rattus_blangorum, Rattus_mindorensis, Rattus_baluensis, Rattus_losea, Rattus_nit.
```

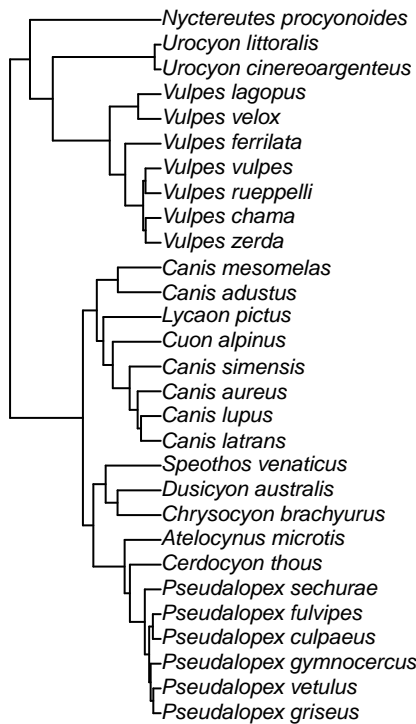
```
##
```

```
## Rooted; includes branch lengths.
```

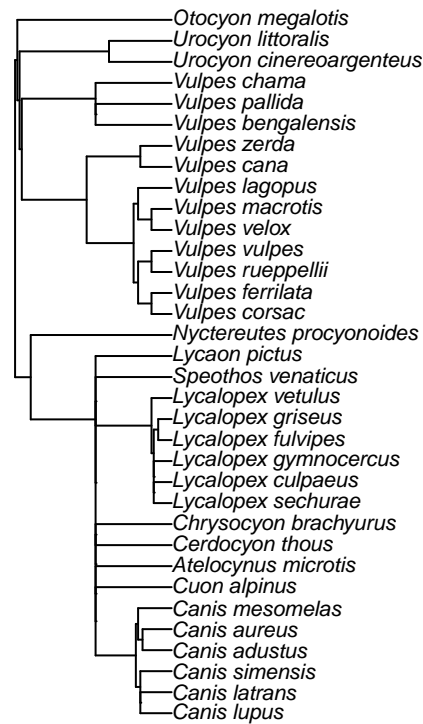
[illegible]

```
# We can also compare the canidae tree we extracted from the mammal megaphylogeny with the one that we
canidae_mrca <- getMRCA(mammal_tree, c("Vulpes_velox", "Lycalopex_griseus"))
Canidae_Tree2 <- extract.clade(mammal_tree, node = canidae_mrca)
par(mfrow = c(1,2))
par(mar = c(1,2,2,0))
plot(Canidae_Tree, cex = 0.7)
title("Canid tree from Carnivoran phylogeny", cex.main = 0.8)
plot(Canidae_Tree2, cex = 0.7)
title("Canid tree from mammal megaphylogeny", cex.main = 0.8)
```

Canid tree from Carnivoran phylogeny



Canid tree from mammal megaphylogeny



2. Plot your tree with tip labels and a scale bar.
3. Annotate at least 2 clades on your tree with some information. Try to find some biologically relevant info for this (e.g., Freshwater/Saltwater, Africa/Asia/Australia, venomous/non-venomous, etc). Tip labels are optional.
4. Plot a PhyloPic image on your tree. You may substitute a different image if your taxon is not in the PhyloPic database. Tip labels are optional here.

Write your tree to a new file. Write a .csv file containing the species in your tree. Now you can open the .csv file as a spreadsheet and begin collecting data for your project!

NOTE As we can see from the example above, the subclade extracted from a megaphylogeny can be very different compared to a subclade extracted from a smaller phylogeny. Therefore, for your final project, it may be better to use a different phylogeny compared to the one you extracted from one of these larger trees, especially if the trees has a lot of polytomies (like the canid example). If you chose a clade from the mammal, bird, or reptile megaphylogenies (or if the tree you're looking at has a lot of polytomies), begin searching the literature to find a better resolved tree.