Supplementary analysis for building phylogeny for Pace of ecology drives temporal perception across the Animal Kingdom

Haarlem, C., Hynes, C,. Jackson, A.L., Healy, K.

17/12/2025

Introduction

This document outlines how the super tree was created by grafting relevant phylogenies together. See the supplementary document for more details on the sources on each of the phylogenies.

Packages

To upload the phylogeny and for other functions associated with handling the phylogeny we used the `phytools`, `caper` and `phangorn` packages. For some more general functions we will use the `devtools` and `MASS` packages. To directly upload some of the phylogenies we will use the `fishtree` and `rotl` and to check synonoms we will use `taxize`

```r
library(phytools)
library(caper)
library(phangorn)

library(devtools)
library(MASS)

library(taxize)
library(fishtree)
library(rotl)
library(phangorn)
```

We will also load some functions that we built specifically for this analysis The phylo_bind_functions script has functions to bind the phylogenies and include error in the nodes joining together trees.

```r
source("phylo_bind_functions.R")
```

Read in data

Next we load in the data so we have the species list to match up the tree at the end.

Read in the phylogenies

Lets now upload each of the phylogenies we are going to use to build our super-dooper tree. We used the Metazoan phylogeny from the Open Tree of Life project (https://ot39.opentreeoflife.org/opentree/argus/opentree10.4@ott93302) as a backbone. For Mammals we used a distribution of 100 trees from Kuhn et al 2011. For Aves we used the Jetz et al 2012 distribution of trees. For lizards we used the Pyron et al 2014 phylogeny.

For our analysis we will use a distribution of 100 supertrees, however as many of the trees will not be suitable as the first appearance from fossil ranges might overlap between nodes resulting in negative distances for certain topologies we will draw 1000 trees from any distribution available. We will first join all the trees together and then at the end trim the tips to just the species in our dataset.

*Please note that as the separate phylogenies are too large to host on Github they will need to be downloaded seperately from there sources. However, we do include the output distribution of trees (Haarlem_et_al_tree.tre) that is the result of this script within the repository.

```
#Metazoan backbone
met_tree <- read.tree("metazoa.tre")
met_tree  <- makeLabel(met_tree)
#Homo_sapiens tip needs to get dropped as it is also present in the FritzTree
#phylogeny and will cause issues.
met_tree <- drop.tip(met_tree,"Homo_sapiens")
met_tree$node.label<-NULL

#Mammal phylogeny
mammal_phy <- read.nexus("FritzTree.rs200k.100trees.tre")

#Aves phylogeny
aves_phy <- read.tree("BirdzillaHackett10.tre")

#Lizard phylogeny
Lizard_phy <- read.tree("liz_and_snake_time_tree.txt")
lizard_phy<-makeLabel(Lizard_phy)
#this removes the node labels which can cause duplicate problems
lizard_phy$node.label<-NULL
#this fixes rounding errors due to format change to make it ultrametric.
lizard_phy <-chronoMPL(lizard_phy)


#Turtle phylogeny
Turt_phy <- read.nexus("Turtle_Posterior_100.tre")

#amphibian phylo
amph_phy <- read.tree("Amphi.tre")
#Again drop the Homo_sapiens tip which was included as an outgroup
amph_phy <- drop.tip(amph_phy, "Homo_sapiens")
#this removes the node lables which can cause duplicate problems
amph_phy$node.label<-NULL
amph_phy <-chronoMPL(amph_phy)

#Shark
shark_phy <- read.nexus("Shark_10.cal.tree.nex")


# lets also set the number of trees we will build.
sample.no <- 1000
```

Join birds and turtles

First we join birds and turtles together. For the common ancestor for turtles and aves we use the stem dates proposed by W. G. Joyce, J. F. Parham, T. R. Lyson, R. C. M. Warnock, P. C. J. Donoghue, A divergence

dating analysis of turtles using fossil calibrations: An example of best practices.J. Paleontol. 87, 612–634 (2013).

```r
diapsid_o <- jiggle.bind(x = aves_phy,
                         y = Turt_phy,
                         sample = 1000,
                         min.age = 156,
                         max.age = 250)


diapsid_nultra <- diapsid_o[[1]]
diapsid_node <- diapsid_o[[2]]


is_ultra_diapsid <- vector()
for(i in 1:length(diapsid_nultra)) {
  is_ultra_diapsid[i] <- is.ultrametric(diapsid_nultra[[i]])
  }

ultra_row_d <- which(is_ultra_diapsid ==TRUE)
diapsid <- diapsid_nultra[ultra_row_d]

class(diapsid) <- "multiPhylo"
```

Add Lepidosauria

To link the Aves_turtles trees to the common ancestor of Lepidosauria we used a range of 259-285 Myr from Jones et al 2013.

```r
length_bird_lizard <- as.numeric(length(diapsid))

bird_lizard_o <- jiggle.bind(x = diapsid,
                         y = lizard_phy,
                         sample = length_bird_lizard,
                         min.age = 259,
                         max.age = 285)


bird_lizard_nultra <- bird_lizard_o[[1]]
bird_lizard_node <- bird_lizard_o[[2]]


is_ultra_bird_lizard<- vector()
for(i in 1:length(bird_lizard_nultra)) {
  is_ultra_bird_lizard[i] <- is.ultrametric(bird_lizard_nultra[[i]])
  }

ultra_row_bl <- which(is_ultra_bird_lizard ==TRUE)
bird_lizard <- bird_lizard_nultra[ultra_row_bl]

class(bird_lizard) <- "multiPhylo"
```

Add Amniotes

For the common ancestor of amniotes, we used the fossil Archerpeton anthracos (Holotype: RM 12056, Author: Carroll 1964, Reisz and Müller, Epoch: Westphalian A Canada Nova Scotia, Age: 318.1 – 314.6 Myr, Dating: International Commission on Stratigraphy 2009).

```r
length_amniote <- as.numeric(length(bird_lizard))

amniote_tree_o <- jiggle.bind(x = bird_lizard,
                              y = mammal_phy,
                              sample = length_amniote,
                              min.age = 314.6,
                              max.age = 318.1)


amniote_tree_nultra <- amniote_tree_o[[1]]
amniote_node <- amniote_tree_o[[2]]


is_ultra_amniote_tree <- vector()
for(i in 1:length(amniote_tree_nultra)) {
  is_ultra_amniote_tree[i] <- is.ultrametric(amniote_tree_nultra[[i]])
  }

ultra_rowam <- which(is_ultra_amniote_tree ==TRUE)
amniote_tree <- amniote_tree_nultra[ultra_rowam]


class(amniote_tree) <- "multiPhylo"
```

Add Amphibians

For the common ancestor between Amphibians and amniotes we used 347-352Mya form Hime et al (2021).

```r
length_amni_amphi <- as.numeric(length(amniote_tree))


amni_amphi_tree_o <- jiggle.bind(x = amniote_tree,
                                 y = amph_phy,
                                 sample = length_amni_amphi,
                                 min.age = 359.2,
                                 max.age = 359.2)


amni_amphi_tree_nultra <- amni_amphi_tree_o[[1]]
amni_amphi_node <- amni_amphi_tree_o[[2]]


is_ultra_amni_amphi_tree <- vector()
for(i in 1:length(amni_amphi_tree_nultra)) {
  is_ultra_amni_amphi_tree[i] <- is.ultrametric(amni_amphi_tree_nultra[[i]])
  }

ultra_rowaa <- which(is_ultra_amni_amphi_tree ==TRUE)
```

```
amni_amphi_tree <- amni_amphi_tree_nultra[ultra_rowaa]


class(amni_amphi_tree) <- "multiPhylo"
```

Add Actinopterygians

To add fish we used the common ancestor of Actinopterygians and Sarcopterygians was dated using the fossil
Guiyu oneiros (Holotype: IVPP V15541, Zhu et al 2013, Epoch: Late Ludlow, Silurian, Kuanti Formation;
Qujing, Yunnan, China., Age: 419 Myr).

Here we use the `fishtree_phylogeny()` function which pulls the phylogeny from the Fish Tree of Life API.
Due to its size we trim the fish tree down to the species we have here.

```
#get the fish species in our dataset
fish_data <- cff_clean[cff_clean$Class %in% c("Actinopterygii"),]
fish_data <- data.frame(species_match1 = unique(fish_data$Species_phylo),
                        species_match = unique(fish_data$Species_phylo))

#retrieve the fish phylogeny
w_fish_tree <- fishtree_phylogeny(type = c("chronogram"))


fish_tree <- comparative.data(phy = w_fish_tree,
                              data = fish_data,
                              names.col = "species_match" ,
                              force.root = TRUE)$phy
length_fish <- as.numeric(length(amni_amphi_tree))

#join the trees together at 219
fish_tetrapod_tree_o <- jiggle.bind(x = amni_amphi_tree,
                                    y = fish_tree,
                                    sample = length_fish,
                                    min.age = 419,
                                    max.age = 419)


fish_tetrapod_tree_nultra <- fish_tetrapod_tree_o[[1]]
fish_tetrapod_node <- fish_tetrapod_tree_o[[2]]


is_ultra_fish_tetrapod_tree <- vector()
for(i in 1:length(fish_tetrapod_tree_nultra)) {
  is_ultra_fish_tetrapod_tree[i] <- is.ultrametric(fish_tetrapod_tree_nultra[[i]])

  fish_tetrapod_tree_nultra[[i]]$node.label = "NA"


  }

ultra_rowff <- which(is_ultra_fish_tetrapod_tree ==TRUE)
fish_tetrapod_tree <- fish_tetrapod_tree_nultra[ultra_rowff]

class(fish_tetrapod_tree) <- "multiPhylo"
```

Add Chondrichthyes

For the common ancestor between Osteichthyans and Chondrichthyes we used the fossil Entelognathus primordialis (Holotype: IVPP V18620, Zhu et al 2009, Epoch: Xiaoxiang Reservoir, Qujing, Yunnan, China Kuanti Formation. Age: 420).

```r
shark_data <- cff_clean[cff_clean$Class %in% c("Elasmobranchii"),]
shark_data <- data.frame(species_match1 = unique(shark_data$Species_phylo),
                         species_match = unique(shark_data$Species_phylo))


length_shark <- as.numeric(length(fish_tetrapod_tree))


shark_tree <- list()
for(i in 1:length_shark){
shark_tree[[i]] <- comparative.data(phy = shark_phy[[i]],
                           data = shark_data,
                           names.col = "species_match" ,
                           force.root = TRUE)$phy

}

class(shark_tree) <- "multiPhylo"

##Need to get this working again

shark_fish_tree_o <- jiggle.bind(x = fish_tetrapod_tree,
                                 y = shark_tree,
                                 sample = length_shark,
                                 min.age = 420,
                                 max.age = 420)


shark_fish_tree_nultra <- shark_fish_tree_o[[1]]
shark_fish_tree_node <- shark_fish_tree_o[[2]]

is_ultra_shark_fish_tree <- vector()
for(i in 1:length(shark_fish_tree_nultra)) {
  is_ultra_shark_fish_tree[i] <- is.ultrametric(shark_fish_tree_nultra[[i]])

  shark_fish_tree_nultra[[i]]$node.label = "NA"

  }

ultra_rowss <- which(is_ultra_shark_fish_tree ==TRUE)
shark_fish_tree <- shark_fish_tree_nultra[ultra_rowss]


class(shark_fish_tree) <- "multiPhylo"
```

Add Agnata

Add in the river lampray we used a divergance divrergance time of 434 ± 8 from Delsuc et al 2018 A phylogenomic framework and timescale for comparative studies of tunicates.

```r
lamp_data <- cff_clean[cff_clean$Species_phylo %in% c("Lampetra_fluviatilis",
                                                      "Mordacia_praecox",
                                                      "Mordacia_mordax"),]

met_tree_l <- met_tree

met_tree_l$node.label <- NA

for(i in 1:(length(lamp_data$Species_phylo))){

  met_tree_l$tip.label[grep(lamp_data$Species_phylo[i],met_tree_l$tip.label)] <- sub("_ott.*", "",met_t

}


lamp_tree <- comparative.data(phy = met_tree_l,
                              data = lamp_data,
                              names.col = "Species_phylo" ,
                              force.root = TRUE)$phy


class(lamp_tree)<-"phylo"
lamp_tree$node.label = "NA"


length_lamp <- as.numeric(length(shark_fish_tree))

chordata_tree_o <- jiggle.bind(x = lamp_tree,
                               y = shark_fish_tree,
                               sample = length_lamp,
                               min.age = 426,
                               max.age = 442)

chordata_tree <- chordata_tree_o[[1]]

for( i in 1:length_lamp){
chordata_tree[[i]]$node.label = "NA"
}

chordata_tree_node <- chordata_tree_o[[2]]
class(chordata_tree) <- "multiPhylo"
```

Add Ambulacraria

For the split between Ambulacraria and Chordata we used the 535-567 Mya range for the common ancestor from Delsuc et al (2018).

```r
echino_tip <-list(edge=matrix(c(2,1),1,2),
          tip.label= "Acanthaster_planci",
```

```
            edge.length=1.0,
            Nnode=1)

class(echino_tip)<-"phylo"

length_echino <- as.numeric(length(chordata_tree))

echino_tree_o <- jiggle.bind(x = echino_tip,
                                y = chordata_tree,
                                sample = length_echino,
                                min.age = 535,
                                max.age = 567)

echino_tree <- echino_tree_o[[1]]

for( i in 1:length_echino){
echino_tree[[i]]$node.label = "NA"
}

echino_tree_node <- echino_tree_o[[2]]
class(echino_tree) <- "multiPhylo"
```

Add Protostomes

For the common ancestor between deuterostomes and protostomes we use the dates of 588-610 Myr and for Bilateria and Cnidaria we used dates of between 553 and 636 Mya from Delsuc et al (2018). Here we use the Open Tree of Life (Hinchliff et al 2015) so we need to clip that tree here to the species in our dataset.

```
invert_data <- cff_clean[cff_clean$Class %in% c("Gastropoda",
                                "Polychaeta",
                                "Cephalopoda",
                                "Insecta",
                                "Malacostraca",
                                "Arachnida"),]

invert_data <- data.frame(Species_phylo = unique(invert_data$Species_phylo),
                    Species_phylo2 = unique(invert_data$Species_phylo))

met_tree2 <- met_tree

met_tree2$node.label <- NA

for(i in 1:(length(invert_data$Species_phylo))){

  met_tree2$tip.label[grep(invert_data$Species_phylo[i],met_tree2$tip.label)] <-
    sub("_ott.*", "",met_tree$tip.label[grep(invert_data$Species_phylo[i],
                                      met_tree$tip.label)])

}


#Remove the duplicates
met_tree_dr <- drop.tip(met_tree2, met_tree2$tip.label[duplicated(met_tree2$tip.label)])
```

```
invert_tree <- comparative.data(phy = met_tree_dr,
                                data = invert_data,
                                names.col = "Species_phylo" ,
                                force.root = TRUE)

invert_tree_ult <-compute.brlen(invert_tree$phy,
                          method = "Grafen",
                          power = 1)

invert_tree_ult$edge.length <- invert_tree_ult $edge.length*587

length_bi <- as.numeric(length(chordata_tree))

bi_lat_tree_o <- jiggle.bind(x = invert_tree_ult,
                                  y = echino_tree,
                                  sample = length_bi,
                                  min.age = 588,
                                  max.age = 610)



bi_lat_tree <- bi_lat_tree_o[[1]]


for( i in 1:length_bi){
bi_lat_tree[[i]]$node.label = "NA"
}


bi_lat_node <- bi_lat_tree_o[[2]]
class(bi_lat_tree) <- "multiPhylo"
```

Add Cnidaria

To add in the jellyfish species we use the Bilateria and Cnidaria divergance dates of between 553 and 636 Mya from Delsuc et al (2018).

```
jelly_tip <-list(edge=matrix(c(2,1),1,2),
          tip.label= "Tripedalia_cystophora",
          edge.length=1.0,
          Nnode=1)

class(jelly_tip)<-"phylo"

length_jelly <- as.numeric(length(bi_lat_tree))

jelly_tree_o <- jiggle.bind(x = jelly_tip,
                                  y = bi_lat_tree,
                                  sample = length_jelly,
                                  min.age = 553,
                                  max.age = 636)

jelly_tree <- jelly_tree_o[[1]]
```

```
for( i in 1:length_jelly){
jelly_tree[[i]]$node.label = "NA"
}

jelly_tree_node <- jelly_tree_o[[2]]
class(jelly_tree) <- "multiPhylo"
```

Final Tree

We now need to clip the final distribution of trees to our data to make sure it all matches up. We also need to check if any trees have negative branch lengths and drop them.

```
final_tree <- list()

cff_clean_match <- data.frame(spec_1 = unique(cff_clean$Species_phylo),
                              spec = unique(cff_clean$Species_phylo))

##Now we clean each of the trees so it only has the species from the dataset in
##it
for(i in 1:(length(jelly_tree))){

  final_tree[[i]] <- comparative.data(phy = jelly_tree[[i]],
                                      data = cff_clean_match,
                                      names.col = "spec",
                                      force.root = TRUE)$phy
}


final_cff_data <- comparative.data(phy = jelly_tree[[1]],
                                   data = cff_clean_match,
                                   names.col = "spec",
                                   force.root = TRUE)


class(final_tree) <- "multiPhylo"


final_tree <- final_tree[is.ultrametric(final_tree)]

final_tree <- final_tree[1:100]


plot(final_tree[[1]],
     type = "fan",
     cex = 0.15)
```
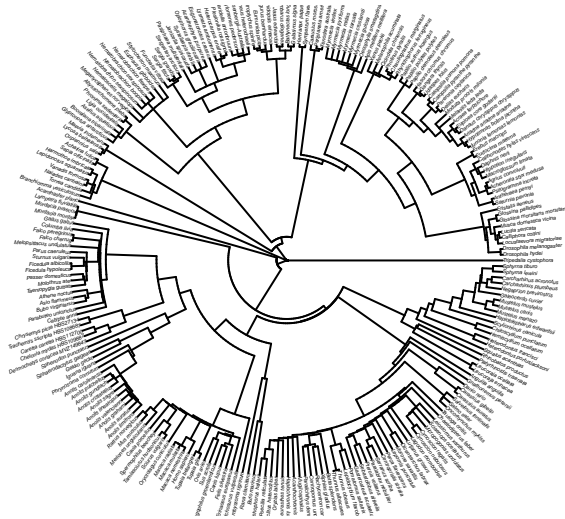
Write tree

After these checks we save the phylogeny

```r
write.tree(final_tree, file = "Haarlem_et_al_tree.tre")
```