Tree and space analysis function

Ty Tuff and Bruno Vilela

1/9/2017

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

## This module analyzes the results from module 1 and returns a list based on how many values each stat returns  
## Ty Tuff and Bruno Vilela  
## 24 August 2016  
  
###### Specify function ##############################  
  
Module\_2 <- function(Module\_1\_output) {  
 cat("\nAnalyzing: 0% [")  
 if (any(is.na(Module\_1\_output))) {  
 cat("----------]")  
 return(NA)  
 } else {  
 if (Ntip(Module\_1\_output$mytree) == 2) {  
 cat("----------]")  
 return(NA)  
 } else {  
  
 this\_tree <- Module\_1\_output$mytree  
 this\_world <- Module\_1\_output$myWorld  
  
 ##### (0) Pull necessary variables from simulated trees and organize into a single object for all the tests below to pull from.  
  
 #str(all\_trees)  
 #str(this\_tree)  
  
 ## 0a) Branch lengths  
 Branch\_Lengths <- this\_tree$edge.length  
 number\_of\_branches <- length(Branch\_Lengths)  
  
 # Anchor test = PD (Faith's phylogenetic diversity)  
 Pylo\_diversity\_is\_sum\_of\_BL <- sum(Branch\_Lengths)  
  
 # avPD -- Average phylogenetic diversity  
 average\_phylogenetic\_diversity\_is\_mean\_of\_BL <- mean(Branch\_Lengths)  
  
 variance\_Pylo\_diversity\_is\_variance\_of\_BL <- var(Branch\_Lengths)  
 cat("-")  
 ## 0b) Pairwise distance between tips  
 Pairwise\_dist <- cophenetic(this\_tree)  
 cat("-")  
 # 2b) Pairwise distance -- Sum of pairwise distances  
  
 # F -- Extensive quadratic entropy  
 F\_quadratic\_entropy\_is\_sum\_of\_PD <- sum(Pairwise\_dist)  
  
 #Mean inter-species distances  
  
 # Anchor test = MPD (mean pairwise distance)  
  
 Mean\_pairwise\_distance <- mean(Pairwise\_dist)  
  
 cat("-")  
 #Pairwise distance/all distances -- Variance of pairwise distances  
  
 # Anchor test = VPD (variation of pairwise distance)  
  
 variance\_pairwise\_distance <- var(as.vector(Pairwise\_dist))  
  
 ## 0c) Phylogenetic isolation  
  
 # Using equal.splits method, faster computation  
 Evolutionary\_distinctiveness <- evol.distinct2(this\_tree, type = "equal.splits")  
 cat("-")  
 # ED - Summed evolutionary distinctiveness  
  
 Evolutionary\_distinctiveness\_sum <- sum(Evolutionary\_distinctiveness)  
  
 ## 3d) Phylogenetic isolation -- Mean of species evolutionary distinctiveness  
  
 # mean(ED)  
  
 mean\_Phylogenetic\_isolation <- mean(Evolutionary\_distinctiveness)  
  
 ## 4d) Phylogenetic isolation -- Variance of species isolation metrics  
  
 #var(ED)  
  
 variance\_Phylogenetic\_isolation <- var(Evolutionary\_distinctiveness)  
 cat("-")  
  
 ## Tree topology  
  
 #Gamma index  
  
 ltts <- ltt(this\_tree, gamma = TRUE, plot = FALSE)  
 lineages\_through\_time <- as.numeric(ltts[[1]])  
 time\_steps <- as.numeric(ltts[[2]])  
 gamma <- ltts[[3]]  
 gamma\_p\_value <- ltts[[4]]  
 cat("-")  
  
 ##### (5) Tree metric -- Macroevolutionary - Rate and rate changes ###############  
 ##################################################  
  
 ## Speciation vs extinction rates and Net diversification  
 bds <- bd(this\_tree)  
 speciation\_rate <- bds[1]  
 extinction\_rate <- bds[2]  
 extinction\_per\_speciation <- bds[3]  
 speciation\_minus\_extinction <- bds[4]  
 cat("-")  
  
 ## Speciation vs extinction rates and Net diversification dependent on trait  
 N.for.dom <- table(this\_world[, 6])  
 if(length(N.for.dom) == 2) {  
 par.div.dep <- DivDep( mytree = this\_tree, myWorld = this\_world)  
 trait\_1\_speciation <- par.div.dep[1]  
 trait\_2\_speciation <- par.div.dep[2]  
 trait\_1\_extinction <- par.div.dep[3]  
 trait\_2\_extinction <- par.div.dep[4]  
 transition\_from\_trait\_1\_to\_2 <- par.div.dep[5]  
 transition\_from\_trait\_2\_to\_1 <- par.div.dep[6]  
 transition\_rate\_ratio\_1to2\_over\_2to1 <- transition\_from\_trait\_1\_to\_2/transition\_from\_trait\_2\_to\_1  
 cat("-")  
  
 ## Crown age per trait AUC and effect size  
 tip.length <- this\_tree$edge.length[this\_tree$edge[, 2] %in% 1:Ntip(this\_tree)]  
 tip.length <- (tip.length - min(tip.length)) / (max(tip.length) - min(tip.length))  
 this\_trait <- this\_world[match(this\_tree$tip.label, this\_world[, 8]), 6]  
 tip.length.2 <- tip.length[this\_trait == 2]  
 tip.length.1 <- tip.length[this\_trait == 1]  
 model <- glm(as.factor(this\_trait) ~ log(tip.length + 1),  
 family = "binomial")  
 effect.size <- model$coefficients[2]  
 plot(y = this\_trait - 1, x= log(tip.length))  
 p <- predict(model, as.factor(this\_trait), type = "resp")  
 points(y = p, x = log(tip.length), col = "red")  
 pr <- prediction(p, as.factor(this\_trait))  
 auc.model <- performance(pr, measure = "auc")@y.values[[1]]  
  
 ## Phylogenetic signal (D)  
 Phylogenetic\_signal <- Dsig(mytree = this\_tree, myWorld = this\_world)  
 cat("-")  
  
 ## Spatial Analysis  
 nbs0 <- knearneigh(as.matrix(this\_world[, 2:3]), k = 7, longlat = TRUE)  
 nbs <- knn2nb(nbs0, sym = TRUE) # 7 symmetric neighbors  
 nbs.listw <- nb2listw(nbs)  
 factors.nbs <- as.factor(ifelse(is.na(this\_world[, 6]), 3, this\_world[, 6]))  
 spatial.tests <- joincount.test(fx = factors.nbs, listw = nbs.listw)  
 spatial.tests.fora <- spatial.tests[[1]]$statistic  
 spatial.tests.dom <- spatial.tests[[2]]$statistic  
 prevalence <- (N.for.dom[1] - N.for.dom[2]) / sum(N.for.dom)  
 cat("-")  
 } else {  
 trait\_1\_speciation <- NA  
 trait\_2\_speciation <- NA  
 trait\_1\_extinction <- NA  
 trait\_2\_extinction <- NA  
 transition\_from\_trait\_1\_to\_2 <- NA  
 transition\_from\_trait\_2\_to\_1 <- NA  
 transition\_rate\_ratio\_1to2\_over\_2to1 <- NA  
 Phylogenetic\_signal <- NA  
 spatial.tests.fora <- NA  
 spatial.tests.dom <- NA  
 auc.model <- NA  
 effect.size <- NA  
 prevalence <- ifelse(names(table(this\_world[, 6])[1]) == "1", 1,  
 -1)  
 cat("---")  
  
 }  
  
 results\_summary\_matrix\_1 <- cbind(  
  
 number\_of\_branches,  
 Pylo\_diversity\_is\_sum\_of\_BL,  
 average\_phylogenetic\_diversity\_is\_mean\_of\_BL,  
 variance\_Pylo\_diversity\_is\_variance\_of\_BL,  
  
 F\_quadratic\_entropy\_is\_sum\_of\_PD,  
 Mean\_pairwise\_distance,  
 variance\_pairwise\_distance,  
  
 Evolutionary\_distinctiveness\_sum,  
 mean\_Phylogenetic\_isolation,  
 variance\_Phylogenetic\_isolation,  
  
 gamma,  
 gamma\_p\_value,  
 speciation\_rate,  
 extinction\_rate,  
 extinction\_per\_speciation,  
 speciation\_minus\_extinction,  
 trait\_1\_speciation,  
 trait\_2\_speciation ,  
 trait\_1\_extinction ,  
 trait\_2\_extinction ,  
 transition\_from\_trait\_1\_to\_2 ,  
 transition\_from\_trait\_2\_to\_1 ,  
 transition\_rate\_ratio\_1to2\_over\_2to1 ,  
 Phylogenetic\_signal,  
 spatial.tests.fora,  
 spatial.tests.dom,  
 prevalence,  
 auc.model,  
 effect.size  
 )  
 rownames(results\_summary\_matrix\_1) <- 1  
  
 results\_summary\_matrix\_2 <- cbind(  
 c(Evolutionary\_distinctiveness,NA),  
 lineages\_through\_time,  
 time\_steps  
 )  
 colnames(results\_summary\_matrix\_2) <- c("Evolutionary\_distinctiveness", "lineages\_through\_time", "time\_steps")  
 head(results\_summary\_matrix\_2)  
  
 ### Returns from function in list form  
 returns <- list(  
 #Branch\_Lengths,  
 #Pairwise\_dist,  
 results\_summary\_matrix\_1,  
 results\_summary\_matrix\_2  
  
 )  
  
 names(returns) <- c(  
 #"Branch\_Lengths",  
 #"Pairwise\_distance",  
 "results\_summary\_of\_single\_value\_outputs",  
 "results\_summary\_matrix\_of\_multi\_value\_outputs"  
 )  
 cat("] 100%")  
  
 return(returns)  
  
 }  
 }  
}  
  
#Module\_2(myOut)

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.