##Create Primate Phylogenetic Trees

install.packages("rotl")

install.packages("ape")

install.packages("phangorn")

library(phangorn)

library(rotl)

library(ape)

##Full Species List Tree

primates <- tnrs\_match\_names(names = "Primates")

print(primates)

primate\_tree <- tol\_subtree(ott\_id = primates$ott\_id)

plot.phylo(primate\_tree)

write.nexus(primate\_tree, file = "primate\_tree.nexus")

prunedspecies\_list <- c("Allenopithecus\_nigroviridis", "Allocebus\_trichotis", "Allochrocebus\_lhoesti", "Allochrocebus\_preussi", "Allochrocebus\_solatus", "Alouatta\_arctoidea", "Alouatta\_belzebul", "Alouatta\_caraya", "Alouatta\_discolor", "Alouatta\_juara", "Alouatta\_nigerrima", "Alouatta\_palliata", "Alouatta\_pigra", "Alouatta\_sara", "Alouatta\_seniculus", "Alouatta\_ululata", "Aotus\_azarae", "Aotus\_brumbacki", "Aotus\_griseimembra", "Aotus\_jorgehernandezi", "Aotus\_lemurinus", "Aotus\_miconax", "Aotus\_nancymaae", "Aotus\_nigriceps", "Aotus\_trivirgatus", "Aotus\_vociferans", "Aotus\_zonalis", "Arctocebus\_aureus", "Arctocebus\_calabarensis", "Ateles\_belzebuth", "Ateles\_fusciceps", "Ateles\_geoffroyi", "Ateles\_marginatus", "Ateles\_paniscus", "Avahi\_betsileo", "Avahi\_cleesei", "Avahi\_laniger", "Avahi\_meridionalis", "Avahi\_mooreorum", "Avahi\_occidentalis", "Avahi\_peyrierasi", "Avahi\_ramanantsoavani", "Avahi\_unicolor", "Brachyteles\_arachnoides", "Brachyteles\_hypoxanthus", "Cacajao\_ayresi", "Cacajao\_calvus", "Cacajao\_hosomi", "Cacajao\_melanocephalus", "Callicebus\_barbarabrownae", "Callicebus\_caquetensis", "Callicebus\_coimbrai", "Callicebus\_discolor", "Callicebus\_dubius", "Callicebus\_melanochir", "Callicebus\_nigrifrons", "Callicebus\_ornatus", "Callicebus\_personatus", "Callimico\_goeldii", "Callithrix\_aurita", "Callithrix\_flaviceps", "Callithrix\_geoffroyi", "Callithrix\_humeralifer", "Callithrix\_humilis", "Callithrix\_jacchus", "Callithrix\_kuhlii", "Callithrix\_penicillata", "Callithrix\_pygmaea", "Carlito\_syrichta", "Cebuella\_niveiventris", "Cebus\_aequatorialis", "Cebus\_albifrons", "Cebus\_brunneus", "Cebus\_capucinus", "Cebus\_cesarae", "Cebus\_cuscinus", "Cebus\_imitator", "Cebus\_kaapori", "Cebus\_leucocephalus", "Cebus\_malitiosus", "Cebus\_olivaceus", "Cebus\_unicolor", "Cebus\_versicolor", "Cebus\_yuracus", "Cephalopachus\_bancanus", "Cercocebus\_agilis", "Cercocebus\_atys", "Cercocebus\_chrysogaster", "Cercocebus\_galeritus", "Cercocebus\_sanjei", "Cercocebus\_torquatus", "Cercopithecus\_albogularis", "Cercopithecus\_ascanius", "Cercopithecus\_buettikoferi", "Cercopithecus\_lowei", "Cercopithecus\_cephus", "Cercopithecus\_denti", "Cercopithecus\_diana", "Cercopithecus\_doggetti", "Cercopithecus\_dryas", "Cercopithecus\_erythrogaster", "Cercopithecus\_hamlyni", "Cercopithecus\_kandti", "Cercopithecus\_lomamiensis", "Cercopithecus\_mitis", "Cercopithecus\_mona",

"Cercopithecus\_neglectus", "Cercopithecus\_nicticans", "Cercopithecus\_petaurista", "Cercopithecus\_pogonias", "Cercopithecus\_roloway", "Cercopithecus\_sclateri", "Cercopithecus\_wolfi", "Cheirogaleus\_andysabini", "Cheirogaleus\_crossleyi", "Cheirogaleus\_lavasoensis", "Cheirogaleus\_major", "Cheirogaleus\_medius", "Cheirogaleus\_minusculus", "Cheirogaleus\_shethi", "Cheirogaleus\_sibreei", "Cheirogaleus\_thomasi", "Cheracebus\_lucifer", "Cheracebus\_lugens", "Cheracebus\_medemi", "Cheracebus\_purinus", "Cheracebus\_regulus", "Cheracebus\_torquatus", "Chiropotes\_albinasus", "Chiropotes\_sagulata", "Chiropotes\_israelita", "Chiropotes\_satanas", "Chiropotes\_utahickae", "Chlorocebus\_aethiops", "Chlorocebus\_cynosuros", "Chlorocebus\_djamdjamensis", "Chlorocebus\_pygerythrus", "Chlorocebus\_sabaeus", "Colobus\_angolensis", "Colobus\_guereza", "Colobus\_polykomos", "Colobus\_satanas", "Colobus\_vellerosus", "Daubentonia\_madagascariensis", "Erythrocebus\_patas", "Erythrocebus\_poliophaeus", "Eulemur\_albifrons", "Eulemur\_cinereiceps", "Eulemur\_coronatus", "Eulemur\_flavifrons", "Eulemur\_fulvus", "Eulemur\_macaco", "Eulemur\_mongoz", "Eulemur\_rubriventer", "Eulemur\_rufifrons", "Eulemur\_rufus", "Eulemur\_sanfordi", "Euoticus\_elegantulus", "Euoticus\_pallidus", "Galago\_alleni", "Galago\_gabonensis", "Galago\_gallarum", "Galago\_granti", "Galago\_matschiei", "Galago\_minor", "Galago\_moholi", "Galago\_senegalensis",

"Galagoides\_kumbirensis", "Galagoides\_orinus", "Galagoides\_rondoensis", "Galagoides\_thomasi", "Galagoides\_zanzibaricus", "Gorilla\_beringei", "Gorilla\_gorilla", "Hapalemur\_alaotrensis", "Hapalemur\_aureus", "Hapalemur\_griseus", "Hapalemur\_occidentalis", "Hoolock\_hoolock", "Hoolock\_leuconedys", "Hoolock\_tianxing", "Hylobates\_agilis", "Hylobates\_albibarbis", "Hylobates\_klossii", "Hylobates\_lar", "Hylobates\_moloch", "Hylobates\_muelleri", "Hylobates\_pileatus", "Indri\_indri", "Lagothrix\_cana", "Lagothrix\_flavicauda", "Lagothrix\_lagotricha", "Lagothrix\_lugens", "Lagothrix\_poeppigii", "Lagothrix\_tschudii", "Lemur\_catta", "Leontopithecus\_caissara", "Leontopithecus\_chrysomelas", "Leontopithecus\_chrysopygus", "Leontopithecus\_rosalia", "Lepilemur\_aeeclis", "Lepilemur\_ahmansonorum", "Lepilemur\_ankaranensis", "Lepilemur\_betsileo", "Lepilemur\_dorsalis", "Lepilemur\_edwardsi", "Lepilemur\_fleuretae",

"Lepilemur\_grewcockorum", "Lepilemur\_hollandorum", "Lepilemur\_hubbardorum", "Lepilemur\_jamesorum", "Lepilemur\_leucopus", "Lepilemur\_manasamody", "Lepilemur\_microdon", "Lepilemur\_milanoii", "Lepilemur\_mitsinjoensis", "Lepilemur\_mittermeieri", "Lepilemur\_mustelinus", "Lepilemur\_otto", "Lepilemur\_petteri", "Lepilemur\_randrianasoloi", "Lepilemur\_ruficaudatus", "Lepilemur\_sahamalazensis", "Lepilemur\_scottorum", "Lepilemur\_seali", "Lepilemur\_septentrionalis", "Lepilemur\_tymerlachsoni", "Lepilemur\_wrightae", "Lophocebus\_albigena", "Lophocebus\_aterrimus", "Loris\_lydekkerianus", "Loris\_tardigradus", "Macaca\_arctoides", "Macaca\_assamensis", "Macaca\_brunnescens", "Macaca\_cyclopis", "Macaca\_fascicularis", "Macaca\_fuscata", "Macaca\_hecki", "Macaca\_leonina", "Macaca\_leucogenys", "Macaca\_maura", "Macaca\_mulatta", "Macaca\_munzala", "Macaca\_nemestrina", "Macaca\_nigra", "Macaca\_nigrescens", "Macaca\_ochreata", "Macaca\_pagensis", "Macaca\_radiata", "Macaca\_siberu", "Macaca\_silenus", "Macaca\_sinica", "Macaca\_speciosa", "Macaca\_sylvanus", "Macaca\_thibetana", "Macaca\_tonkeana", "Mandrillus\_leucophaeus", "Mandrillus\_sphinx", "Mico\_acariensis", "Mico\_argentatus", "Mico\_chrysoleucus", "Mico\_emiliae", "Mico\_humeralifer", "Mico\_intermedius", "Mico\_leucippe", "Mico\_marcai", "Mico\_mauesi", "Mico\_melanurus", "Mico\_munduruku", "Mico\_nigriceps", "Mico\_rondoni", "Mico\_saterei", "Microcebus\_arnholdi", "Microcebus\_berthae", "Microcebus\_bongolavensis", "Microcebus\_boraha", "Microcebus\_danfossi", "Microcebus\_ganzhorni", "Microcebus\_gerpi", "Microcebus\_griseorufus", "Microcebus\_jollyae", "Microcebus\_lehilahytsara", "Microcebus\_macarthurii", "Microcebus\_mamiratra", "Microcebus\_manitatra", "Microcebus\_margotmarshae", "Microcebus\_marohita", "Microcebus\_mittermeieri", "Microcebus\_murinus", "Microcebus\_myoxinus", "Microcebus\_ravelobensis", "Microcebus\_rufus", "Microcebus\_sambiranensis", "Microcebus\_simmonsi", "Microcebus\_tanosi", "Microcebus\_tavaratra", "Miopithecus\_ogouensis", "Miopithecus\_talapoin", "Mirza\_coquereli", "Mirza\_zaza", "Nasalis\_larvatus", "Nomascus\_annamensis", "Nomascus\_concolor", "Nomascus\_gabriellae", "Nomascus\_hainanus", "Nomascus\_leucogenys", "Nomascus\_nasutus", "Nomascus\_siki", "Nycticebus\_bancanus", "Nycticebus\_bengalensis", "Nycticebus\_borneanus", "Nycticebus\_coucang", "Nycticebus\_javanicus",

"Nycticebus\_kayan", "Nycticebus\_pygmaeus", "Otolemur\_crassicaudatus", "Otolemur\_garnettii", "Otolemur\_monteiri", "Pan\_paniscus", "Pan\_troglodytes", "Papio\_anubis", "Papio\_cynocephalus", "Papio\_hamadryas", "Papio\_kindae", "Papio\_papio", "Papio\_ursinus", "Paragalago\_cocos", "Paragalago\_zanzibaricus", "Perodicticus\_potto", "Phaner\_electromontis", "Phaner\_furcifer", "Phaner\_pallescens", "Phaner\_parienti", "Piliocolobus\_badius", "Piliocolobus\_bouvieri", "Piliocolobus\_epieni", "Piliocolobus\_foai", "Piliocolobus\_gordonorum", "Piliocolobus\_kirkii", "Piliocolobus\_lulindicus", "Piliocolobus\_oustaleti", "Piliocolobus\_parmentieri", "Piliocolobus\_pennantii", "Piliocolobus\_preussi", "Piliocolobus\_rufomitratus", "Piliocolobus\_semlikiensis", "Piliocolobus\_tephrosceles", "Piliocolobus\_tholloni", "Piliocolobus\_waldronae", "Pithecia\_aequatorialis", "Pithecia\_albicans", "Pithecia\_cazuzai", "Pithecia\_chrysocephala", "Pithecia\_hirsuta", "Pithecia\_inusta", "Pithecia\_irrorata", "Pithecia\_isabela", "Pithecia\_milleri", "Pithecia\_mittermeieri", "Pithecia\_monachus", "Pithecia\_napensis", "Pithecia\_pissinattii", "Pithecia\_pithecia", "Pithecia\_rylandsi", "Pithecia\_vanzolinii", "Plecturocebus\_aureipalatii", "Plecturocebus\_baptista", "Plecturocebus\_bernhardi", "Plecturocebus\_brunneus", "Plecturocebus\_cinerascens", "Plecturocebus\_cupreus", "Plecturocebus\_donacophilus", "Plecturocebus\_grovesi", "Plecturocebus\_hoffmannsi", "Plecturocebus\_miltoni", "Plecturocebus\_modestus", "Plecturocebus\_moloch", "Plecturocebus\_oenanthe", "Plecturocebus\_olallae", "Plecturocebus\_pallescens", "Plecturocebus\_stephennashi", "Plecturocebus\_toppini", "Plecturocebus\_urubambensis", "Plecturocebus\_vieirai", "Pongo\_abelii", "Pongo\_pygmaeus", "Pongo\_tapanuliensis", "Presbytis\_canicrus", "Presbytis\_chrysomelas", "Presbytis\_comata", "Presbytis\_femoralis", "Presbytis\_frontata", "Presbytis\_hosei", "Presbytis\_melalophos", "Presbytis\_natunae", "Presbytis\_pileata", "Presbytis\_potenziani", "Presbytis\_robinsoni", "Presbytis\_rubicunda", "Presbytis\_sabana", "Presbytis\_siamensis", "Presbytis\_thomasi", "Procolobus\_verus", "Prolemur\_simus", "Propithecus\_candidus", "Propithecus\_coquereli", "Propithecus\_deckenii", "Propithecus\_diadema", "Propithecus\_edwardsi", "Propithecus\_perrieri", "Propithecus\_tattersalli", "Propithecus\_verreauxi",

"Pseudopotto\_martini", "Pygathrix\_cinerea", "Pygathrix\_nemaeus", "Pygathrix\_nigripes", "Rhinopithecus\_avunculus", "Rhinopithecus\_bieti", "Rhinopithecus\_brelichi", "Rhinopithecus\_roxellana", "Rhinopithecus\_strykeri", "Rungwecebus\_kipunji", "Saguinus\_bicolor", "Saguinus\_cruzlimai", "Saguinus\_fuscicollis", "Saguinus\_geoffroyi", "Saguinus\_graellsi", "Saguinus\_imperator", "Saguinus\_inustus", "Saguinus\_labiatus", "Saguinus\_leucopus", "Saguinus\_martinsi", "Saguinus\_melanoleucus", "Saguinus\_midas", "Saguinus\_mystax", "Saguinus\_niger", "Saguinus\_nigricollis", "Saguinus\_oedipus", "Saguinus\_tripartitus", "Saguinus\_ursulus", "Saimiri\_boliviensis", "Saimiri\_cassiquiarensis", "Saimiri\_oerstedii", "Saimiri\_sciureus", "Saimiri\_ustus", "Saimiri\_vanzolinii", "Sapajus\_apella", "Sapajus\_cay", "Sapajus\_flavius", "Sapajus\_libidinosus", "Sapajus\_nigritus", "Sapajus\_xanthosternos", "Sciurocheirus\_makandensis", "Semnopithecus\_ajax", "Semnopithecus\_entellus", "Semnopithecus\_hector", "Semnopithecus\_hypoleucos", "Semnopithecus\_natunae", "Semnopithecus\_priam", "Semnopithecus\_schistaceus", "Simias\_concolor", "Symphalangus\_syndactylus", "Tarsius\_dentatus", "Tarsius\_fuscus", "Tarsius\_lariang", "Tarsius\_pelengensis", "Tarsius\_pumilus", "Tarsius\_sangirensis", "Tarsius\_spectrumgurskyae", "Tarsius\_supriatnai", "Tarsius\_tarsier", "Tarsius\_tumpara", "Tarsius\_wallacei", "Theropithecus\_gelada", "Trachypithecus\_auratus", "Trachypithecus\_barbei", "Trachypithecus\_cristatus", "Trachypithecus\_delacouri", "Trachypithecus\_ebenus", "Trachypithecus\_francoisi", "Trachypithecus\_geei", "Trachypithecus\_germaini", "Trachypithecus\_hatinhensis", "Trachypithecus\_johnii", "Trachypithecus\_laotum", "Trachypithecus\_mauritius", "Trachypithecus\_obscurus", "Trachypithecus\_phayrei", "Trachypithecus\_poliocephalus", "Trachypithecus\_selangorensis", "Trachypithecus\_shortridgei", "Trachypithecus\_vetulus", "Varecia\_rubra", "Varecia\_variegata")

prunedspecies\_match <- tnrs\_match\_names(names = prunedspecies\_list)

prunedspecies\_tree <- tol\_induced\_subtree(ott\_ids = prunedspecies\_match$ott\_id)

prunedspecies\_tree <- compute.brlen(prunedspecies\_tree, method = "Grafen")

plot.phylo(prunedspecies\_tree)

write.nexus(prunedspecies\_tree, file = "pruned\_species\_tree.nexus")

species\_list <- c("Ateles\_fusciceps", "Ateles\_geoffroyi", "Callithrix\_jacchus", "Callithrix\_penicillata",

"Cebus\_albifrons", "Cebus\_capucinus", "Cebus\_imitator", "Cebus\_olivaceus",

"Cercocebus\_atys", "Cercopithecus\_ascanius", "Cercopithecus\_mona",

"Chlorocebus\_pygerythrus", "Colobus\_guereza", "Colobus\_vellerosus", "Erythrocebus\_patas",

"Eulemur\_fulvus", "Galago\_moholi", "Gorilla\_beringei", "Gorilla\_gorilla", "Hylobates\_lar",

"Lagothrix\_lagotricha", "Lagothrix\_poeppigii", "Leontopithecus\_rosalia", "Lophocebus\_albigena",

"Macaca\_arctoides", "Macaca\_fascicularis", "Macaca\_fuscata", "Macaca\_maura", "Macaca\_mulatta",

"Macaca\_nemestrina", "Macaca\_nigra", "Macaca\_radiata", "Macaca\_silenus", "Macaca\_sylvanus",

"Macaca\_thibetana", "Macaca\_tonkeana", "Miopithecus\_talapoin", "Nasalis\_larvatus", "Pan\_paniscus",

"Pan\_troglodytes", "Papio\_anubis", "Papio\_cynocephalus", "Papio\_hamadryas", "Pongo\_abelii",

"Pongo\_pygmaeus", "Propithecus\_verreauxi", "Rhinopithecus\_roxellana", "Saguinus\_fuscicollis",

"Saguinus\_geoffroyi", "Saguinus\_mystax", "Saguinus\_oedipus", "Saimiri\_sciureus", "Sapajus\_apella",

"Sapajus\_libidinosus", "Sapajus\_nigritus", "Sapajus\_xanthosternos", "Semnopithecus\_entellus",

"Symphalangus\_syndactylus", "Theropithecus\_gelada", "Trachypithecus\_johnii")

species\_match <- tnrs\_match\_names(names = species\_list)

species\_tree <- tol\_induced\_subtree(ott\_ids = species\_match$ott\_id)

if (is.null(species\_tree$edge.length)) {

species\_tree$edge.length <- rep(1, nrow(species\_tree$edge))

}

plot.phylo(species\_tree)

write.nexus(species\_tree, file = "SSBspecies\_tree.nexus")

##Filtered SSB Species Tree

species\_list <- c("Callithrix\_jacchus", "Colobus\_vellerosus", "Gorilla\_beringei", "Gorilla\_gorilla",

"Hylobates\_lar", "Macaca\_arctoides", "Macaca\_fascicularis", "Macaca\_fuscata",

"Macaca\_mulatta", "Macaca\_nemestrina", "Macaca\_nigra", "Macaca\_sylvanus",

"Macaca\_thibetana", "Macaca\_tonkeana", "Nasalis\_larvatus", "Pan\_paniscus",

"Pan\_troglodytes", "Papio\_anubis", "Papio\_cynocephalus", "Rhinopithecus\_roxellana",

"Saimiri\_sciureus", "Sapajus\_nigritus", "Semnopithecus\_entellus")

species\_match <- tnrs\_match\_names(names = species\_list)

species\_tree <- tol\_induced\_subtree(ott\_ids = species\_match$ott\_id)

if (is.null(species\_tree$edge.length)) {

species\_tree$edge.length <- rep(1, nrow(species\_tree$edge))

}

plot.phylo(species\_tree)

write.nexus(species\_tree, file = "filteredSSBspecies\_tree.nexus")

##################################################

##Phylogenetic Tree Figure

install.packages("BiocManager")

BiocManager::install("ggtree")

install.packages(c("ggplot2", "remotes", "stringr"))

remotes::install\_github("YuLab-SMU/ggtreeExtra")

library(ape)

library(dplyr)

library(stringr)

library(ggplot2)

library(ggtree)

full\_tree <- read.nexus("pruned\_species\_tree.nexus")

p <- ggtree(full\_tree, layout = "circular") +

theme(

legend.position = "none",

plot.margin = margin(2, 2, 2, 2, "cm")

)

p

##Phylogenetic Tree (genus level)

install.packages("BiocManager")

BiocManager::install("ggtree")

install.packages(c("ggplot2", "remotes", "stringr"))

remotes::install\_github("YuLab-SMU/ggtreeExtra")

library(ape)

library(dplyr)

library(stringr)

library(ggplot2)

library(ggtree)

full\_tree <- read.nexus("pruned\_species\_tree.nexus")

species\_to\_genus <- tibble(

species = full\_tree$tip.label,

genus = word(full\_tree$tip.label, 1, sep = "\_")

)

keep\_tips <- species\_to\_genus %>%

distinct(genus, .keep\_all = TRUE) %>%

pull(species)

genus\_tree <- keep.tip(full\_tree, keep\_tips)

genus\_tree$tip.label <- word(genus\_tree$tip.label, 1, sep = "\_")

p <- ggtree(genus\_tree, layout = "circular") +

geom\_tiplab(size = 6, offset = 0.15) +

theme(

legend.position = "none",

plot.margin = margin(2, 2, 2, 2, "cm")

)

p

####################################################################################

##Phylogenetic Signal of SSB

library(ape)

library(geiger)

tree <- read.nexus("pruned\_species\_tree.nexus")

ssb\_data <- read.csv("full\_dataframe.csv", stringsAsFactors = FALSE)

ssb\_data <- ssb\_data[, c("nexus\_name", "ssb\_presence")]

ssb\_data$nexus\_name <- as.character(ssb\_data$nexus\_name)

ssb\_data$ssb\_presence <- as.numeric(ssb\_data$ssb\_presence)

tree <- drop.tip(tree, setdiff(tree$tip.label, ssb\_data$nexus\_name))

ssb\_data <- ssb\_data[ssb\_data$nexus\_name %in% tree$tip.label, ]

tree <- multi2di(tree)

ssb\_vector <- ssb\_data$ssb\_presence

names(ssb\_vector) <- ssb\_data$nexus\_name

ssb\_vector <- ssb\_vector + 1

fit <- fitDiscrete(tree, ssb\_vector, model = "ER", transform = "lambda")

print(fit)

##################################################################################

##Phylogenetic Regression Models

install.packages("caper")

install.packages("ape")

install.packages("phylolm")

install.packages("ggplot2")

install.packages("dplyr")

library(caper)

library(ape)

library(phylolm)

library(ggplot2)

library(dplyr)

#Load your data

data <- read.csv("full\_dataframe.csv")

ssbdata <- read.csv("filtered\_ssb\_dataframe.csv")

tree <- read.nexus("pruned\_species\_tree.nexus")

ssbtree <- read.nexus("filteredSSBspecies\_tree.nexus")

#Check tip names

rownames(data) <- data$nexus\_name

pruned\_tree <- drop.tip(tree, setdiff(tree$tip.label, rownames(data)))

rownames(ssbdata) <- ssbdata$nexus\_name

pruned\_ssb\_tree <- drop.tip(ssbtree, setdiff(ssbtree$tip.label, rownames(ssbdata)))

#Prevalence of male and female SSB

ssbdata$logMale <- ifelse(ssbdata$male\_ssb==0, 0, log(ssbdata$male\_ssb/ssbdata$sampling\_hours))

ssbdata$logFemale <- ifelse(ssbdata$female\_ssb==0, 0, log(ssbdata$female\_ssb/ssbdata$sampling\_hours))

##Temperature

#Temperature and SSB Presence

model1 <- phyloglm(ssb\_presence ~ mean\_temp\_2019, data = data, phy = pruned\_tree, method = "logistic\_MPLE")

summary(model1)

#Temperature and SSB Prevalence

bm\_model2 <- phylolm(log(ssb\_prevalence) ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model2)

pagel\_model3 <- phylolm(log(ssb\_prevalence) ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model3)

#Temperature and Male SSB

bm\_model4 <- phylolm(logMale ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model4)

pagel\_model5 <- phylolm(logMale ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model5)

#Temperature and Female SSB

bm\_model6 <- phylolm(logFemale ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model6)

pagel\_model7 <- phylolm(logFemale ~ mean\_temp\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model7)

##Precipitation

#Precipitation and SSB Presence

model8 <- phyloglm(ssb\_presence ~ mean\_precip\_2019, data = data, phy = pruned\_tree, method = "logistic\_MPLE")

summary(model8)

#Precipitation and SSB Prevalence

bm\_model9 <- phylolm(log(ssb\_prevalence) ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model9)

pagel\_model10 <- phylolm(log(ssb\_prevalence) ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model10)

#Precipitation and Male SSB

bm\_model11 <- phylolm(logMale ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model11)

pagel\_model12 <- phylolm(logMale ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model12)

#Precipitation and Female SSB

bm\_model13 <- phylolm(logFemale ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model13)

pagel\_model14 <- phylolm(logFemale ~ mean\_precip\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model14)

##Cloud Cover

#Cloud Cover and SSB Presence

model15 <- phyloglm(ssb\_presence ~ mean\_cloud\_2019, data = data, phy = pruned\_tree, method = "logistic\_MPLE")

summary(model15)

#Cloud Cover and SSB Prevalence

bm\_model16 <- phylolm(log(ssb\_prevalence) ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model16)

pagel\_model17 <- phylolm(log(ssb\_prevalence) ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model17)

#Cloud Cover and Male SSB

bm\_model18 <- phylolm(logMale ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model18)

pagel\_model19 <- phylolm(logMale ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model19)

#cloud Cover and Female SSB

bm\_model20 <- phylolm(logFemale ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model20)

pagel\_model21 <- phylolm(logFemale ~ mean\_cloud\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model21)

##Resource Availability (Biomass)

#Biomass and SSB Presence

model22 <- phyloglm(ssb\_presence ~ mean\_biomass\_2019, data = data, phy = pruned\_tree, method = "logistic\_MPLE", log.alpha.bound = 5)

summary(model22)

#Biomass and SSB Prevalence

bm\_model23 <- phylolm(log(ssb\_prevalence) ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model23)

pagel\_model24 <- phylolm(log(ssb\_prevalence) ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model24)

#Biomass and Male SSB

bm\_model25 <- phylolm(logMale ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model25)

pagel\_model26 <- phylolm(logMale ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model26)

#Biomass and Female SSB

bm\_model27 <- phylolm(logFemale ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model27)

pagel\_model28 <- phylolm(logFemale ~ mean\_biomass\_2019, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model28)

##Predation Pressure

#Predation and SSB Presence

model29 <- phyloglm(ssb\_presence ~ scale(predator\_count), data = data, phy = pruned\_tree, method = "logistic\_MPLE", btol = 5, log.alpha.bound = 6)

summary(model29)

#Predation and SSB Prevalence

bm\_model30 <- phylolm(log(ssb\_prevalence) ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model30)

pagel\_model31 <- phylolm(log(ssb\_prevalence) ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model31)

#Predation and Male SSB

bm\_model32 <- phylolm(logMale ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model32)

pagel\_model33 <- phylolm(logMale ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model33)

#Predation and Female SSB

bm\_model34 <- phylolm(logFemale ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model34)

pagel\_model35 <- phylolm(logFemale ~ scale(predator\_count), data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model35)

##Sexual Size Dimorphism

#Sexual Dimorphism and SSB Presence

model36 <- phyloglm(ssb\_presence ~ size\_dimorphism\_presence, data = data, phy = pruned\_tree, method = "logistic\_MPLE", log.alpha.bound = 6)

summary(model36)

##Median Body Size

#Median Body Size and SSB Presence

data$median\_body\_size <- as.numeric(data$median\_body\_size)

model37 <- phyloglm(ssb\_presence ~ log(median\_body\_size), data = data, phy = pruned\_tree, method = "logistic\_MPLE")

summary(model37)

#Median Body Size and SSB Prevalence

bm\_model38 <- phylolm(log(ssb\_prevalence) ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model38)

pagel\_model39 <- phylolm(log(ssb\_prevalence) ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model39)

#Median Body Size and Male SSB

bm\_model40 <- phylolm(logMale ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model40)

pagel\_model41 <- phylolm(logMale ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model41)

#Median Body Size and Female SSB

bm\_model42 <- phylolm(logFemale ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model42)

pagel\_model43 <- phylolm(logFemale ~ median\_body\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model43)

##Body Size Range

#Body Size Range and SSB Presence

model44 <- phyloglm(ssb\_presence ~ log(body\_size\_range +1), data = data, phy = pruned\_tree, method = "logistic\_MPLE", log.alpha.bound = 6)

summary(model44)

#Body Size Range and SSB Prevalence

bm\_model45 <- phylolm(log(ssb\_prevalence) ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model45)

pagel\_model46 <- phylolm(log(ssb\_prevalence) ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model46)

#Body Size Range and Male SSB

bm\_model47 <- phylolm(logMale ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model47)

pagel\_model48 <- phylolm(logMale ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model48)

#Body Size Range and Female SSB

bm\_model49 <- phylolm(logFemale ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model49)

pagel\_model50 <- phylolm(logFemale ~ body\_size\_range, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model50)

##Lifespan

#Lifespan and SSB Presence

model51 <- phyloglm(ssb\_presence ~ log(lifespan +1), data = data, phy = pruned\_tree, method = "logistic\_MPLE", log.alpha.bound = 6)

summary(model51)

#Lifespan and SSB Prevalence

bm\_model52 <- phylolm(log(ssb\_prevalence) ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model52)

pagel\_model53 <- phylolm(log(ssb\_prevalence) ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model53)

#Lifespan and Male SSB

bm\_model54 <- phylolm(logMale ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model54)

pagel\_model55 <- phylolm(logMale ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model55)

#Lifespan and Female SSB

bm\_model56 <- phylolm(logFemale ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model56)

pagel\_model57 <- phylolm(logFemale ~ lifespan, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model57)

##Sex Ratio

#Sex Ratio and SSB Presence

model58 <- phyloglm(ssb\_presence ~ sex\_ratio, data = data, phy = pruned\_tree, method = "logistic\_MPLE", btol=10, log.alpha.bound = 6)

summary(model58)

#Sex Ratio and SSB Prevalence

bm\_model59 <- phylolm(log(ssb\_prevalence) ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model59)

pagel\_model60 <- phylolm(log(ssb\_prevalence) ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model60)

#Sex Ratio and Male SSB

bm\_model61 <- phylolm(logMale ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree)

summary(model61)

pagel\_model62 <- phylolm(logMale ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model62)

#Sex Ratio and Female SSB

bm\_model63 <- phylolm(logFemale ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model63)

pagel\_model64 <- phylolm(logFemale ~ sex\_ratio, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model64)

##Group Size

#Group Size and SSB Presence

model65 <- phyloglm(ssb\_presence ~ log(median\_group\_size), data = data, phy = pruned\_tree, method = "logistic\_MPLE", log.alpha.bound = 5)

summary(model65)

#Group Size and SSB Prevalence

bm\_model66 <- phylolm(log(ssb\_prevalence) ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model66)

pagel\_model67 <- phylolm(log(ssb\_prevalence) ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model67)

#Group Size and Male SSB

bm\_model68 <- phylolm(logMale ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model68)

pagel\_model69 <- phylolm(logMale ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model69)

#Group Size and Female SSB

bm\_model70 <- phylolm(logFemale ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model70)

pagel\_model71 <- phylolm(logFemale ~ median\_group\_size, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model71)

##Group Structure

#Group Structure and SSB Presence

model72 <- phyloglm(ssb\_presence ~ group\_structure, data = data, phy = pruned\_tree, method = "logistic\_MPLE")

confint(model72)

summary(model72)

#Group Structure and SSB Prevalence

bm\_model73 <- phylolm(log(ssb\_prevalence) ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model73)

pagel\_model74 <- phylolm(log(ssb\_prevalence) ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model74)

#Group Structure and Male SSB

bm\_model75 <- phylolm(logMale ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model75)

pagel\_model76 <- phylolm(logMale ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model76)

#Group Structure and Female SSB

bm\_model77 <- phylolm(logFemale ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model77)

pagel\_model78 <- phylolm(logFemale ~ group\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model78)

##Hierarchical Structure

#Hierarchical Structure and SSB Presence

model79 <- phyloglm(ssb\_presence ~ hierarchical\_structure, data = data, phy = pruned\_tree, method = "logistic\_MPLE", btol=19)

summary(model79)

#Hierarchical Structure and SSB Prevalence

bm\_model80 <- phylolm(log(ssb\_prevalence) ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model80)

pagel\_model81 <- phylolm(log(ssb\_prevalence) ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model81)

#Hierarchical Structure and Male SSB

bm\_model82 <- phylolm(logMale ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model82)

pagel\_model83 <- phylolm(logMale ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model83)

#Hierarchical Structure and Female SSB

bm\_model84 <- phylolm(logFemale ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model84)

pagel\_model85 <- phylolm(logFemale ~ hierarchical\_structure, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model85)

##Mating System

#Mating System and SSB Presence

model86 <- phyloglm(ssb\_presence ~ mating\_system, data = data, phy = pruned\_tree, method = "logistic\_MPLE", btol=16)

summary(model86)

#Mating System and SSB Prevalence

bm\_model87 <- phylolm(log(ssb\_prevalence) ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model87)

pagel\_model88 <- phylolm(log(ssb\_prevalence) ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model88)

#Mating System and Male SSB

bm\_model89 <- phylolm(logMale ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model89)

pagel\_model90 <- phylolm(logMale ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model90)

#Mating System and Female SSB

bm\_model91 <- phylolm(logFemale ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model91)

pagel\_model92 <- phylolm(logFemale ~ mating\_system, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model92)

##Infant Rearing Strategy

#Infant Rearing Strategy and SSB Presence

model93 <- phyloglm(ssb\_presence ~ infant\_rearing\_strategy, data = data, phy = pruned\_tree, method = "logistic\_MPLE", btol = 22, log.alpha.bound = 6)

summary(model93)

#Infant Rearing STrategy and SSB Prevalence

bm\_model94 <- phylolm(log(ssb\_prevalence) ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model94)

pagel\_model95 <- phylolm(log(ssb\_prevalence) ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model95)

#Infant Rearing Strategy and Male SSB

bm\_model96 <- phylolm(logMale ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model96)

pagel\_model97 <- phylolm(logMale ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model97)

#Infant Rearing Strategy and Female SSB

bm\_model98 <- phylolm(logFemale ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree)

summary(bm\_model98)

pagel\_model99 <- phylolm(logFemale ~ infant\_rearing\_strategy, data = ssbdata, phy = pruned\_ssb\_tree, model = "lambda", lower.bound = 0.001, upper.bound = 0.999)

summary(pagel\_model99)

#############################################

##Models for all SSB species (59 species)

#Load your data

all\_ssb\_data <- read.csv("all\_ssb\_species\_dataframe.csv")

all\_ssb\_tree <- read.nexus("SSBspecies\_tree.nexus")

all\_ssb\_data <- all\_ssb\_data[all\_ssb\_data$nexus\_name != "" & !is.na(all\_ssb\_data$nexus\_name), ]

all\_ssb\_data <- all\_ssb\_data[!duplicated(all\_ssb\_data$nexus\_name), ]

#Check tip names

rownames(all\_ssb\_data) <- all\_ssb\_data$nexus\_name

pruned\_ssb\_tree <- drop.tip(all\_ssb\_tree, setdiff(all\_ssb\_tree$tip.label, rownames(all\_ssb\_data)))

#Temperature and all SSB species

ssb\_temp <- phylolm(log(ssb\_prevalence) ~ mean\_temp\_2019, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_temp)

#Precipitation and all SSB species

ssb\_precip <- phylolm(log(ssb\_prevalence) ~ mean\_precip\_2019, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_precip)

#Cloud Cover and all SSB species

ssb\_cloud <- phylolm(log(ssb\_prevalence) ~ mean\_cloud\_2019, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_cloud)

#Biomass and all SSB species

ssb\_biomass <- phylolm(log(ssb\_prevalence) ~ mean\_biomass\_2019, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_biomass)

#Predation and all SSB species

ssb\_predation <- phylolm(log(ssb\_prevalence) ~ scale(predator\_count), data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_predation)

#Median Body Size and all SSB species

ssb\_median\_body\_size <- phylolm(log(ssb\_prevalence) ~ median\_body\_size, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_median\_body\_size)

#Body Size Range and all SSB species

ssb\_body\_size\_range <- phylolm(log(ssb\_prevalence) ~ body\_size\_range, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_body\_size\_range)

#Lifespan and all SSB species

ssb\_lifespan <- phylolm(log(ssb\_prevalence) ~ lifespan, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_lifespan)

#Sex Ratio and all SSB species

bm\_model59 <- phylolm(log(ssb\_prevalence) ~ sex\_ratio, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(bm\_model59)

#Group Size and all SSB species

ssb\_group\_size <- phylolm(log(ssb\_prevalence) ~ median\_group\_size, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_group\_size)

#Group Structure and all SSB species

ssb\_group\_structure <- phylolm(log(ssb\_prevalence) ~ group\_structure, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_group\_structure)

#Hierarchical Structure and all SSB species

ssb\_hierarchy <- phylolm(log(ssb\_prevalence) ~ hierarchical\_structure, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_hierarchy)

#Mating System and all SSB species

ssb\_mating\_system <- phylolm(log(ssb\_prevalence) ~ mating\_system, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_mating\_system)

#Infant Rearing Strategy and all SSB species

ssb\_infant\_rearing <- phylolm(log(ssb\_prevalence) ~ infant\_rearing\_strategy, data = all\_ssb\_data, phy = pruned\_ssb\_tree)

summary(ssb\_infant\_rearing)

#############################################

##Phylogenetic Regression Plots

##Plot Temperature and SSB Presence

library(sf)

library(viridis)

library(rnaturalearth)

library(terra)

grib\_file <- "C:/Users/hyrat/OneDrive - Imperial College London/SSB\_climate/data/global\_climate\_data\_detailed/data.grib"

ssb\_filtered\_file <- "Z:/Chloe Coxshall/Literature Analysis (C1)/Data and Analysis ONLY/filtered\_ssb\_dataframe.csv"

shapefile\_dir <- "C:/Users/hyrat/OneDrive - Imperial College London/SSB\_climate/shapefiles/"

grib\_data <- rast(grib\_file)

temp\_layers <- grib\_data[[grep("2 metre temperature \\[C\\]", names(grib\_data))]]

time\_steps <- time(temp\_layers)

selected\_layers <- which(as.Date(time\_steps) >= as.Date("2019-01-01") &

as.Date(time\_steps) <= as.Date("2019-12-31"))

temperature\_2019 <- temp\_layers[[selected\_layers]]

mean\_temp <- mean(temperature\_2019, na.rm = TRUE)

mean\_temp <- mean\_temp - 273.15

mean\_temp <- project(mean\_temp, "EPSG:4326")

ssb\_filtered <- read.csv(ssb\_filtered\_file)

ssb\_species\_names <- ssb\_filtered$species\_name

zip\_files <- list.files(shapefile\_dir, pattern = "\\.zip$", full.names = TRUE)

ssb\_zips <- zip\_files[tools::file\_path\_sans\_ext(basename(zip\_files)) %in% ssb\_species\_names]

pb <- txtProgressBar(min = 0, max = length(ssb\_zips), style = 3)

shapefiles\_list <- list()

for (i in seq\_along(ssb\_zips)) {

zip\_file <- ssb\_zips[i]

temp\_dir <- tempdir()

unzip(zip\_file, exdir = temp\_dir)

shp\_file <- list.files(temp\_dir, pattern = "\\.shp$", full.names = TRUE)

if (length(shp\_file) > 0) {

sf\_obj <- st\_read(shp\_file[1], quiet = TRUE)

species\_name <- tools::file\_path\_sans\_ext(basename(zip\_file))

sf\_obj$Species.Name <- species\_name

shapefiles\_list[[species\_name]] <- sf\_obj

}

setTxtProgressBar(pb, i)

}

close(pb)

ssb\_ranges <- do.call(rbind, shapefiles\_list)

ggplot() +

geom\_raster(

data = as.data.frame(mean\_temp, xy = TRUE),

aes(x = x, y = y, fill = mean)

) +

geom\_sf(data = world\_map, fill = NA, color = "white", linewidth = 0.2) +

geom\_sf(data = ssb\_ranges, fill = "#009999", color = NA, alpha = 0.6) +

scale\_fill\_viridis\_c(

name = "Temperature (°C)",

option = "magma", direction = -1,

na.value = "grey90"

) +

theme\_minimal(base\_size = 16) +

theme(

panel.grid = element\_blank(),

axis.text = element\_blank(),

axis.ticks = element\_blank(),

axis.title = element\_blank(),

legend.position = "right"

)

##Plot Precipitation and SSB Presence

coefs <- coef(model8)

vcov\_matrix <- vcov(model8)

intercept <- coefs["(Intercept)"]

slope <- coefs["mean\_precip\_2019"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["mean\_precip\_2019", "mean\_precip\_2019"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "mean\_precip\_2019"]

x\_vals <- seq(min(data$mean\_precip\_2019, na.rm = TRUE),

max(data$mean\_precip\_2019, na.rm = TRUE),

length.out = 100)

log\_odds <- intercept + slope \* x\_vals

log\_odds\_var <- var\_intercept +

(x\_vals^2) \* var\_slope +

2 \* x\_vals \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

mean\_precip\_2019 = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = mean\_precip\_2019, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Mean Annual Precipitation (m) for 2019",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Cloud Cover and SSB Presence

coefs <- coef(model15)

vcov\_matrix <- vcov(model15)

intercept <- coefs["(Intercept)"]

slope <- coefs["mean\_cloud\_2019"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["mean\_cloud\_2019", "mean\_cloud\_2019"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "mean\_cloud\_2019"]

x\_vals <- seq(min(data$mean\_cloud\_2019, na.rm = TRUE),

max(data$mean\_cloud\_2019, na.rm = TRUE),

length.out = 100)

log\_odds <- intercept + slope \* x\_vals

log\_odds\_var <- var\_intercept +

(x\_vals^2) \* var\_slope +

2 \* x\_vals \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

mean\_cloud\_2019 = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = mean\_cloud\_2019, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Mean Annual Cloud Cover (%) for 2019",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Biomass and SSB Presence

coefs <- coef(model22)

vcov\_matrix <- vcov(model22)

intercept <- coefs["(Intercept)"]

slope <- coefs["mean\_biomass\_2019"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["mean\_biomass\_2019", "mean\_biomass\_2019"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "mean\_biomass\_2019"]

x\_vals <- seq(min(data$mean\_biomass\_2019, na.rm = TRUE),

max(data$mean\_biomass\_2019, na.rm = TRUE),

length.out = 100)

log\_odds <- intercept + slope \* x\_vals

log\_odds\_var <- var\_intercept +

(x\_vals^2) \* var\_slope +

2 \* x\_vals \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

mean\_biomass\_2019 = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = mean\_biomass\_2019, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Mean Leaf Biomass (2019)",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Predation Pressure and SSB Presence

coefs <- coef(model29)

vcov\_matrix <- vcov(model29)

intercept <- coefs["(Intercept)"]

slope <- coefs["scale(predator\_count)"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["scale(predator\_count)", "scale(predator\_count)"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "scale(predator\_count)"]

x\_vals <- seq(min(data$predator\_count, na.rm = TRUE),

max(data$predator\_count, na.rm = TRUE),

length.out = 100)

scaled\_x <- scale(x\_vals, center = attr(scale(data$predator\_count), "scaled:center"),

scale = attr(scale(data$predator\_count), "scaled:scale"))

log\_odds <- intercept + slope \* as.numeric(scaled\_x)

log\_odds\_var <- var\_intercept +

(as.numeric(scaled\_x)^2) \* var\_slope +

2 \* as.numeric(scaled\_x) \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

predator\_count = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = predator\_count, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Number of Predator Species per Primate Range",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Sexual Dimorphism and SSB Presence

coefs <- coef(model36)

vcov\_matrix <- vcov(model36)

groups <- c("Absent", "Present")

results <- data.frame(

Group = groups,

Mean = NA,

LowerCI = NA,

UpperCI = NA

)

z\_value <- qnorm(0.975)

for (i in seq\_along(groups)) {

if (i == 1) {

log\_odds <- coefs["(Intercept)"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"]

} else {

log\_odds <- coefs["(Intercept)"] + coefs["size\_dimorphism\_presence"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"] +

vcov\_matrix["size\_dimorphism\_presence", "size\_dimorphism\_presence"] +

2 \* vcov\_matrix["(Intercept)", "size\_dimorphism\_presence"]

}

lower\_log\_odds <- log\_odds - z\_value \* sqrt(variance)

upper\_log\_odds <- log\_odds + z\_value \* sqrt(variance)

results[i, "Mean"] <- plogis(log\_odds)

results[i, "LowerCI"] <- plogis(lower\_log\_odds)

results[i, "UpperCI"] <- plogis(upper\_log\_odds)

}

results$ErrorBarType <- ifelse(results$Group == "Present", "solid", "dotted")

my\_colors <- c(

"Absent" = "#009999", # Teal

"Present" = "#800080" # Purple

)

ggplot(results, aes(x = Group, y = Mean)) +

geom\_point(aes(color = Group), size = 5) +

geom\_errorbar(

aes(ymin = LowerCI, ymax = UpperCI, linetype = ErrorBarType),

width = 0.2,

linewidth = 1

) +

labs(y = "Predicted Probability of SSB", x = "Sexual Dimorphism") +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

scale\_color\_manual(values = my\_colors) +

scale\_linetype\_manual(values = c("solid" = "solid", "dotted" = "dotted")) +

guides(color = "none", shape = "none", linetype = "none") +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Median Body Size and SSB Presence

coefs <- coef(model37)

vcov\_matrix <- vcov(model37)

intercept <- coefs["(Intercept)"]

slope <- coefs["log(median\_body\_size)"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["log(median\_body\_size)", "log(median\_body\_size)"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "log(median\_body\_size)"]

x\_vals <- seq(min(log(data$median\_body\_size), na.rm = TRUE),

max(log(data$median\_body\_size), na.rm = TRUE),

length.out = 100)

log\_odds <- intercept + slope \* x\_vals

log\_odds\_var <- var\_intercept +

(x\_vals^2) \* var\_slope +

2 \* x\_vals \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

median\_body\_size = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = median\_body\_size, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Median Adult Body Size (kg)",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Body Size Range and SSB Presence

coefs <- coef(model44)

vcov\_matrix <- vcov(model44)

intercept <- coefs["(Intercept)"]

slope <- coefs["log(body\_size\_range + 1)"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["log(body\_size\_range + 1)", "log(body\_size\_range + 1)"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "log(body\_size\_range + 1)"]

x\_vals <- seq(min(data$body\_size\_range, na.rm = TRUE),

max(data$body\_size\_range, na.rm = TRUE),

length.out = 100)

log\_x <- log(x\_vals + 1)

log\_odds <- intercept + slope \* log\_x

log\_odds\_var <- var\_intercept +

(log\_x^2) \* var\_slope +

2 \* log\_x \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

body\_size\_range = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = body\_size\_range, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Adult Body Size Range (kg)",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Lifespan and SSB Presence

coefs <- coef(model51)

vcov\_matrix <- vcov(model51)

intercept <- coefs["(Intercept)"]

slope <- coefs["log(lifespan + 1)"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["log(lifespan + 1)", "log(lifespan + 1)"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "log(lifespan + 1)"]

x\_vals <- seq(min(data$lifespan, na.rm = TRUE),

max(data$lifespan, na.rm = TRUE),

length.out = 100)

log\_x <- log(x\_vals + 1)

log\_odds <- intercept + slope \* log\_x

log\_odds\_var <- var\_intercept +

(log\_x^2) \* var\_slope +

2 \* log\_x \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

lifespan = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = lifespan, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Lifespan (Years)",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Sex Ratio and SSB Presence

coefs <- coef(model58)

vcov\_matrix <- vcov(model58)

# Extract intercept and slope

intercept <- coefs["(Intercept)"]

slope <- coefs["sex\_ratio"]

# Extract variance and covariance

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["sex\_ratio", "sex\_ratio"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "sex\_ratio"]

# Generate x values (untransformed)

x\_vals <- seq(min(data$sex\_ratio, na.rm = TRUE),

max(data$sex\_ratio, na.rm = TRUE),

length.out = 100)

# Compute predicted log-odds and variance

log\_odds <- intercept + slope \* x\_vals

log\_odds\_var <- var\_intercept +

(x\_vals^2) \* var\_slope +

2 \* x\_vals \* cov\_int\_slope

# Convert to probability scale with 95% CIs

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

# Build prediction dataframe

pred\_df <- data.frame(

sex\_ratio = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

# Plot

ggplot(pred\_df, aes(x = sex\_ratio, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Adult Sex Ratio (Male : Female)",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Group Size

coefs <- coef(model65)

vcov\_matrix <- vcov(model65)

intercept <- coefs["(Intercept)"]

slope <- coefs["log(median\_group\_size)"]

var\_intercept <- vcov\_matrix["(Intercept)", "(Intercept)"]

var\_slope <- vcov\_matrix["log(median\_group\_size)", "log(median\_group\_size)"]

cov\_int\_slope <- vcov\_matrix["(Intercept)", "log(median\_group\_size)"]

x\_vals <- seq(min(data$median\_group\_size, na.rm = TRUE),

max(data$median\_group\_size, na.rm = TRUE),

length.out = 100)

log\_x <- log(x\_vals)

log\_odds <- intercept + slope \* log\_x

log\_odds\_var <- var\_intercept +

(log\_x^2) \* var\_slope +

2 \* log\_x \* cov\_int\_slope

z <- qnorm(0.975)

lower\_log\_odds <- log\_odds - z \* sqrt(log\_odds\_var)

upper\_log\_odds <- log\_odds + z \* sqrt(log\_odds\_var)

pred\_df <- data.frame(

median\_group\_size = x\_vals,

Mean = plogis(log\_odds),

LowerCI = plogis(lower\_log\_odds),

UpperCI = plogis(upper\_log\_odds)

)

ggplot(pred\_df, aes(x = median\_group\_size, y = Mean)) +

geom\_line(color = "#800080", linewidth = 1.2) +

geom\_ribbon(aes(ymin = LowerCI, ymax = UpperCI),

fill = "#800080", alpha = 0.2) +

labs(

x = "Median Group Size",

y = "Predicted Probability of SSB Presence"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Group Structure and SSB Presence

coefs <- coef(model72)

vcov\_matrix <- vcov(model72)

groups <- names(coefs)

groups <- c("Fission-Fusion", groups[-1])

results <- data.frame(

Group = groups,

Mean = NA,

LowerCI = NA,

UpperCI = NA

)

z\_value <- qnorm(0.975)

for (i in seq\_along(groups)) {

if (i == 1) {

log\_odds <- coefs["(Intercept)"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"]

} else {

log\_odds <- coefs["(Intercept)"] + coefs[groups[i]]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"] +

vcov\_matrix[groups[i], groups[i]] +

2 \* vcov\_matrix["(Intercept)", groups[i]]

}

lower\_log\_odds <- log\_odds - z\_value \* sqrt(variance)

upper\_log\_odds <- log\_odds + z\_value \* sqrt(variance)

results[i, "Mean"] <- plogis(log\_odds)

results[i, "LowerCI"] <- plogis(lower\_log\_odds)

results[i, "UpperCI"] <- plogis(upper\_log\_odds)

}

results$Group <- c(

"Fission–Fusion",

"Pair-Living",

"MMMF",

"SMMF",

"Solitary"

)

results$ErrorBarType <- ifelse(

results$Group %in% c("Pair-Living", "Solitary"),

"solid", "dotted"

)

my\_colors <- c(

"Fission–Fusion" = "#800080",

"Pair-Living" = "#BFA2DB",

"MMMF" = "#009999",

"SMMF" = "#0000FF",

"Solitary" = "#66C2A5"

)

ggplot(results, aes(x = Group, y = Mean)) +

geom\_point(aes(color = Group), size = 5) +

geom\_errorbar(

aes(ymin = LowerCI, ymax = UpperCI, linetype = ErrorBarType),

width = 0.2,

linewidth = 1

) +

labs(y = "Predicted Probability of SSB", x = "Group Structure Type") +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25)) +

scale\_color\_manual(values = my\_colors) +

scale\_linetype\_manual(values = c("solid" = "solid", "dotted" = "dotted")) +

guides(color = "none", shape = "none", linetype = "none") +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Hierarchical Structure and SSB Presence

coefs <- coef(model79)

vcov\_matrix <- vcov(model79)

groups <- names(coefs)

groups <- c("Dominant Pair", groups[-1])

results <- data.frame(

Group = groups,

Mean = NA,

LowerCI = NA,

UpperCI = NA

)

z\_value <- qnorm(0.975)

for (i in seq\_along(groups)) {

if (i == 1) {

log\_odds <- coefs["(Intercept)"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"]

} else {

log\_odds <- coefs["(Intercept)"] + coefs[groups[i]]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"] +

vcov\_matrix[groups[i], groups[i]] +

2 \* vcov\_matrix["(Intercept)", groups[i]]

}

lower\_log\_odds <- log\_odds - z\_value \* sqrt(variance)

upper\_log\_odds <- log\_odds + z\_value \* sqrt(variance)

mean\_prob <- plogis(log\_odds)

lower\_prob <- plogis(lower\_log\_odds)

upper\_prob <- plogis(upper\_log\_odds)

results[i, "Mean"] <- mean\_prob

results[i, "LowerCI"] <- lower\_prob

results[i, "UpperCI"] <- upper\_prob

}

results$Group<- c("Dominant Pair", "Egalitarian", "Matriarchal",

"Patriarchal")

results$ErrorBarType <- ifelse(results$Group == "Egalitarian", "dotted", "solid")

ggplot(results, aes(x = Group, y = Mean)) +

geom\_point(aes(color = Group), size = 5) +

geom\_errorbar(

mapping = aes(ymin = LowerCI, ymax = UpperCI, linetype = ErrorBarType),

width = 0.2,

linewidth = 1

) +

labs(y = "Predicted Probability of SSB", x = "Hierarchical Structure Type") +

theme\_bw() +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20)

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25)) +

scale\_color\_manual(values = c(

"Dominant Pair" = "#800080",

"Egalitarian" = "#009999",

"Matriarchal" = "#BFA2DB",

"Patriarchal" = "#0000FF"

)) +

scale\_linetype\_manual(values = c("solid" = "solid", "dotted" = "dotted")) +

guides(color = "none", shape = "none", linetype = "none") +

theme\_minimal(base\_size = 20) +

theme(

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Mating System and SSB Presence

coefs <- coef(model86)

vcov\_matrix <- vcov(model86)

levels <- c("Monogamous",

"mating\_systemPolyandrous",

"mating\_systemPolygynandrous",

"mating\_systemPolygynous")

results <- data.frame(

MatingSystem = c("Monogamous", "Polyandrous", "Polygynandrous", "Polygynous"),

Mean = NA,

LowerCI = NA,

UpperCI = NA

)

z\_value <- qnorm(0.975)

for (i in seq\_along(levels)) {

if (i == 1) {

log\_odds <- coefs["(Intercept)"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"]

} else {

term <- levels[i]

log\_odds <- coefs["(Intercept)"] + coefs[term]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"] +

vcov\_matrix[term, term] +

2 \* vcov\_matrix["(Intercept)", term]

}

lower\_log\_odds <- log\_odds - z\_value \* sqrt(variance)

upper\_log\_odds <- log\_odds + z\_value \* sqrt(variance)

results[i, "Mean"] <- plogis(log\_odds)

results[i, "LowerCI"] <- plogis(lower\_log\_odds)

results[i, "UpperCI"] <- plogis(upper\_log\_odds)

}

results$ErrorBarType <- ifelse(

results$MatingSystem %in% c("Monogamous", "Polygynandrous"),

"solid", "dotted"

)

my\_colors <- c(

"Monogamous" = "#800080",

"Polyandrous" = "#66C2A5",

"Polygynandrous" = "#009999",

"Polygynous" = "#BFA2DB"

)

ggplot(results, aes(x = MatingSystem, y = Mean)) +

geom\_point(aes(color = MatingSystem), size = 5) +

geom\_errorbar(

aes(ymin = LowerCI, ymax = UpperCI, linetype = ErrorBarType),

width = 0.2,

linewidth = 1

) +

labs(

y = "Predicted Probability of SSB Presence",

x = "Mating System"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

scale\_color\_manual(values = my\_colors) +

scale\_linetype\_manual(values = c("solid" = "solid", "dotted" = "dotted")) +

guides(color = "none", linetype = "none") +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

##Plot Infant Rearing Strategy and SSB Presence

coefs <- coef(model93)

vcov\_matrix <- vcov(model93)

levels <- c("allo-parenting",

"infant\_rearing\_strategyMaternal",

"infant\_rearing\_strategyPaternal",

"infant\_rearing\_strategyShared parental care")

results <- data.frame(

Strategy = c("Allo-parenting", "Maternal", "Paternal", "Shared Parental"),

Mean = NA,

LowerCI = NA,

UpperCI = NA

)

z\_value <- qnorm(0.975)

for (i in seq\_along(levels)) {

if (i == 1) {

log\_odds <- coefs["(Intercept)"]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"]

} else {

term <- levels[i]

log\_odds <- coefs["(Intercept)"] + coefs[term]

variance <- vcov\_matrix["(Intercept)", "(Intercept)"] +

vcov\_matrix[term, term] +

2 \* vcov\_matrix["(Intercept)", term]

}

lower\_log\_odds <- log\_odds - z\_value \* sqrt(variance)

upper\_log\_odds <- log\_odds + z\_value \* sqrt(variance)

results[i, "Mean"] <- plogis(log\_odds)

results[i, "LowerCI"] <- plogis(lower\_log\_odds)

results[i, "UpperCI"] <- plogis(upper\_log\_odds)

}

results$ErrorBarType <- ifelse(

results$Strategy %in% c("Allo-parenting", "Maternal"),

"solid", "dotted"

)

my\_colors <- c(

"Allo-parenting" = "#800080",

"Maternal" = "#66C2A5",

"Paternal" = "#009999",

"Shared Parental" = "#BFA2DB"

)

ggplot(results, aes(x = Strategy, y = Mean)) +

geom\_point(aes(color = Strategy), size = 5) +

geom\_errorbar(

aes(ymin = LowerCI, ymax = UpperCI, linetype = ErrorBarType),

width = 0.2,

linewidth = 1

) +

labs(

y = "Predicted Probability of SSB Presence",

x = "Infant Rearing Strategy"

) +

scale\_y\_continuous(breaks = seq(0, 1, by = 0.25), limits = c(0, 1)) +

scale\_color\_manual(values = my\_colors) +

scale\_linetype\_manual(values = c("solid" = "solid", "dotted" = "dotted")) +

guides(color = "none", linetype = "none") +

theme\_minimal(base\_size = 20) +

theme(

axis.title = element\_text(size = 22),

axis.text = element\_text(size = 20),

panel.grid.major.x = element\_blank(),

panel.grid.minor = element\_blank()

)

#####################################

##Structural Equation Model

## 1. Load & Check Data

# Load required packages

install.packages("ape")

install.packages("geiger")

install.packages("phytools")

install.packages("dplyr")

install.packages("lavaan")

install.packages("semPlot")

install.packages("phylolm")

install.packages("lavaanPlot")

install.packages("semTools")

install.packages("tidySEM")

install.packages("psych")

install.packages("GPArotation")

library(ape)

library(geiger)

library(phytools)

library(dplyr)

library(lavaan)

library(semPlot)

library(phylolm)

library(lavaanPlot)

library(semTools)

library(tidySEM)

library(psych)

library(GPArotation)

df <- read.csv("full\_dataframe.csv")

tree <- read.nexus("pruned\_species\_tree.nexus")

name.check(tree, df$nexus\_name)

df <- df[match(tree$tip.label, df$nexus\_name), ]

df <- df[!is.na(df$nexus\_name), ]

rownames(df) <- df$nexus\_name

phylo\_cor <- vcv(tree, corr = TRUE)

traits <- c(

"mean\_temp\_2019" = "temp\_resid",

"mean\_precip\_2019" = "precip\_resid",

"mean\_cloud\_2019" = "cloud\_resid",

"mean\_biomass\_2019" = "biomass\_resid",

"predator\_count" = "predation\_resid",

"lifespan" = "lifespan\_resid",

"size\_dimorphism\_presence" = "sdimorphism\_resid",

"median\_body\_size" = "bodysize\_resid",

"body\_size\_range" = "sizerange\_resid",

"mating\_system" = "matingsystem\_resid",

"infant\_rearing\_strategy" = "infant\_resid",

"group\_structure" = "groupstructure\_resid",

"median\_group\_size" = "groupsize\_resid",

"hierarchical\_structure" = "hierarchy\_resid",

"sex\_ratio" = "sexratio\_resid",

"ssb\_presence" = "SSB\_resid"

)

df$`median\_body\_size` <- as.numeric(df$`median\_body\_size`)

unique(df$`median\_body\_size`[!grepl("^\\d+(\\.\\d+)?$", df$`median\_body\_size`)])

df$`median\_body\_size` <- as.numeric(as.character(df$`median\_body\_size`))

df$mating\_system <- as.integer(as.factor(df$mating\_system))

df$infant\_rearing\_strategy <- as.integer(as.factor(df$infant\_rearing\_strategy))

df$group\_structure <- as.integer(as.factor(df$group\_structure))

df$hierarchical\_structure <- as.integer(as.factor(df$hierarchical\_structure))

run\_pgls\_resid <- function(formula, df, phy, resid\_name) {

trait\_var <- all.vars(formula)[1]

df\_trait <- df[!is.na(df[[trait\_var]]), ]

phy\_pruned <- drop.tip(phy, setdiff(phy$tip.label, df\_trait$nexus\_name))

tryCatch({

model <- phylolm(formula, data = df\_trait, phy = phy\_pruned, model = "BM")

df\_trait[[resid\_name]] <- resid(model)

df\_trait[, c("nexus\_name", resid\_name)]

}, error = function(e) {

message(paste("Skipping", resid\_name, "due to error:", e$message))

return(NULL)

})

}

resid\_df\_list <- list()

for (trait in names(traits)) {

resid\_name <- traits[[trait]]

formula <- as.formula(paste(trait, "~ 1"))

df\_out <- run\_pgls\_resid(formula, df, tree, resid\_name)

if (!is.null(df\_out)) {

resid\_df\_list[[resid\_name]] <- df\_out

}

}

df\_resid <- Reduce(function(x, y) full\_join(x, y, by = "nexus\_name"), resid\_df\_list)

df <- left\_join(df, df\_resid, by = "nexus\_name")

resid\_vars <- grep("\_resid$", names(df), value = TRUE)

data <- df[, resid\_vars]

data <- na.omit(data)

describe(data)

corre <- cor(data)

cor.plot(corre)

eigen(corre)

scree(corre, factors = FALSE)

fa.parallel(data)

fa\_result5 <- fa(data, nfactors = 5, scores = "tenBerge")

fa\_result5

print(fa\_result5, sort = TRUE)

print(loadings(fa\_result5), cutoff = .3, digits = 2)

diagram(fa\_result5)

relevant\_columns <- c("temp\_resid", "precip\_resid", "cloud\_resid", "biomass\_resid",

"predation\_resid", "lifespan\_resid", "sdimorphism\_resid",

"bodysize\_resid", "sizerange\_resid", "matingsystem\_resid",

"infant\_resid", "groupstructure\_resid", "groupsize\_resid",

"hierarchy\_resid", "sexratio\_resid", "SSB\_resid") # Add all relevant columns used in FA

rows\_used\_in\_FA <- complete.cases(df[, relevant\_columns])

df\_used\_in\_FA <- df[rows\_used\_in\_FA, ]

dim(df\_used\_in\_FA)

dim(fa\_result5$scores)

df\_used\_in\_FA <- cbind(df\_used\_in\_FA, fa\_result5$scores)

write.table(df\_used\_in\_FA, file = "datascores.csv", row.names = FALSE)

df\_scaled <- df\_used\_in\_FA

numeric\_vars <- sapply(df\_scaled, is.numeric)

df\_scaled[numeric\_vars] <- scale(df\_scaled[numeric\_vars])

model\_hierarchical <- '

# SSB presence is the observed outcome of latent SSB

SSB\_latent =~ SSB\_resid

# Latent SSB is predicted by social structure

SSB\_latent ~ groupstructure\_resid + matingsystem\_resid + infant\_resid +

hierarchy\_resid + groupsize\_resid

# Social structure is shaped by life history

groupstructure\_resid ~ lifespan\_resid + sdimorphism\_resid + bodysize\_resid + sizerange\_resid

matingsystem\_resid ~ lifespan\_resid + sdimorphism\_resid + bodysize\_resid + sizerange\_resid

infant\_resid ~ lifespan\_resid + sdimorphism\_resid + bodysize\_resid + sizerange\_resid

hierarchy\_resid ~ lifespan\_resid + sdimorphism\_resid + bodysize\_resid + sizerange\_resid

groupsize\_resid ~ lifespan\_resid + sdimorphism\_resid + bodysize\_resid + sizerange\_resid

# Life history is influenced by environmental factors

lifespan\_resid ~ temp\_resid + precip\_resid + cloud\_resid + biomass\_resid + predation\_resid

sdimorphism\_resid ~ temp\_resid + precip\_resid + cloud\_resid + biomass\_resid + predation\_resid

bodysize\_resid ~ temp\_resid + precip\_resid + cloud\_resid + biomass\_resid + predation\_resid

sizerange\_resid ~ temp\_resid + precip\_resid + cloud\_resid + biomass\_resid + predation\_resid

'

fit\_path\_analysis <- sem(model\_hierarchical, data = df\_scaled, estimator = "ML")

summary(fit\_path\_analysis, standardized = TRUE, fit.measures = TRUE)

group\_list <- list(

"Environment" = c("temp\_resid", "precip\_resid", "cloud\_resid", "biomass\_resid", "predation\_resd"),

"Life History" = c("lifespan\_resid", "sdimorphism\_resid", "bodysize\_resid", "sizerange\_resid"),

"Social Structure"= c("groupstructure\_resid", "matingsystem\_resid", "infant\_resid", "hierarchy\_resid", "groupsize\_resid"),

"Latent SSB" = c("SSB\_latent"),

"SSB Presence" = c("SSB\_resid")

)

semPaths(fit\_path\_analysis,

what = "std", # Show standardized estimates

layout = "tree", # Tree layout

whatLabels = "std", # Label with standardized values

style = "lisrel", # Neat SEM style

nCharNodes = 0, # Show full variable names

sizeMan = 6, # Size of observed nodes

sizeLat = 8, # Size of latent node

edge.label.cex = 1.1, # Edge label size

groups = group\_list, # Apply layered grouping

color = list(lat = "skyblue", man = "white"),

borders = TRUE,

residuals = FALSE,

title = FALSE)

model\_trimmed <- '

# Measurement model: latent SSB

SSB\_latent =~ SSB\_resid

# Social structure → latent SSB

SSB\_latent ~ groupstructure\_resid + matingsystem\_resid + infant\_resid

# Life history → social structure

groupstructure\_resid ~ lifespan\_resid + sdimorphism\_resid + sizerange\_resid

matingsystem\_resid ~ sdimorphism\_resid

hierarchy\_resid ~ sdimorphism\_resid

groupsize\_resid ~ lifespan\_resid + sdimorphism\_resid

# Environment → life history

lifespan\_resid ~ temp\_resid + precip\_resid + predation\_resid

sdimorphism\_resid ~ precip\_resid

'

fit\_trimmed <- sem(model\_trimmed, data = df\_scaled, estimator = "ML")

summary(fit\_trimmed, standardized = TRUE, fit.measures = TRUE)

group\_list\_trimmed <- list(

"Environment" = c("temp\_resid", "precip\_resid", "predation\_resd"),

"Life History" = c("lifespan\_resid", "sdimorphism\_resid", "sizerange\_resid"),

"Social Structure"= c("groupstructure\_resid", "matingsystem\_resid", "infant\_resid", "hierarchy\_resid", "groupsize\_resid"),

"Latent SSB" = c("SSB\_latent"),

"SSB Presence" = c("SSB\_resid")

)

semPaths(fit\_trimmed,

what = "std",

layout = "tree",

whatLabels = "std",

style = "lisrel",

groups = group\_list\_trimmed,

title = FALSE,

sizeMan = 6, sizeLat = 8,

edge.label.cex = 1.1)

###########################################################

##Bayesian Alternative SEM

install.packages("ape")

install.packages("geiger")

install.packages("dplyr")

install.packages("phylolm")

install.packages("Rphylopars")

install.packages("brms")

install.packages("cmdstanr", repos = c("https://mc-stan.org/r-packages/", getOption("repos")))

cmdstanr::check\_cmdstan\_toolchain(fix = TRUE)

cmdstanr::check\_cmdstan\_toolchain()

cmdstanr::install\_cmdstan()

library(ape)

library(geiger)

library(dplyr)

library(phylolm)

library(Rphylopars)

library(brms)

df <- read.csv("full\_dataframe.csv")

tree <- read.nexus("pruned\_species\_tree.nexus")

colnames(df) <- make.names(colnames(df))

rownames(df) <- df$Nexus.Name

name.check(tree, df)

tree <- drop.tip(tree, "Galagoides\_zanzibaricus\_ott552575")

df <- df[!(rownames(df) == "Galagoides\_demidoffi\_ott7655192"), ]

df <- df[df$nexus\_name %in% tree$tip.label, ]

rownames(df) <- df$nexus\_name

tree <- drop.tip(tree, setdiff(tree$tip.label, df$nexus\_name))

traits <- c(

"mean\_temp\_2019" = "temp\_resid",

"mean\_precip\_2019" = "precip\_resid",

"mean\_cloud\_2019" = "cloud\_resid",

"mean\_biomass\_2019" = "biomass\_resid",

"predator\_count" = "predation\_resid",

"lifespan" = "lifespan\_resid",

"size\_dimorphism\_presence" = "sdimorphism\_resid",

"median\_body\_size" = "bodysize\_resid",

"body\_size\_range" = "sizerange\_resid",

"mating\_system" = "matingsystem\_resid",

"infant\_rearing\_strategy" = "infant\_resid",

"group\_structure" = "groupstructure\_resid",

"median\_group\_size" = "groupsize\_resid",

"hierarchical\_structure" = "hierarchy\_resid",

"sex\_ratio" = "sexratio\_resid"

)

resid\_df\_list <- list()

for (trait in names(traits)) {

resid\_name <- traits[[trait]]

try({

model <- phylolm(as.formula(paste(trait, "~ 1")), data = df, phy = tree, model = "BM")

resid\_values <- resid(model)

out <- data.frame(nexus\_name = names(resid\_values))

out[[resid\_name]] <- resid\_values

resid\_df\_list[[resid\_name]] <- out

}, silent = TRUE)

}

names(resid\_df\_list)

df\_resid <- Reduce(function(x, y) full\_join(x, y, by = "nexus\_name"), resid\_df\_list)

df <- left\_join(df, df\_resid, by = "nexus\_name")

write.csv(df, "residualized\_dataframe.csv", row.names = FALSE)

tree <- drop.tip(tree, setdiff(tree$tip.label, df$nexus\_name))

df <- df[df$nexus\_name %in% tree$tip.label, ]

rownames(df) <- df$nexus\_name

cont\_vars <- grep("\_resid$", names(df), value = TRUE)

trait\_data <- df[, c("nexus\_name", cont\_vars)]

colnames(trait\_data)[1] <- "species"

stopifnot(all(tree$tip.label %in% trait\_data$species))

stopifnot(all(trait\_data$species %in% tree$tip.label))

pp\_fit <- phylopars(

trait\_data = trait\_data,

tree = tree,

model = "BM"

)

cont\_pred <- pp\_fit$anc\_recon[ tree$tip.label , cont\_vars ]

colnames(cont\_pred) <- paste0(cont\_vars, "\_pred")

df <- cbind(df, cont\_pred)

df$SSB\_binary <- ifelse(!is.na(df$ssb\_presence), 1, 0)

A <- ape::vcv.phylo(tree, corr = TRUE)

options(brms.backend = "cmdstanr")

ssb\_fit <- brm(

SSB\_binary ~ 1 + (1 | gr(nexus\_name, cov = A)),

data = df,

family = bernoulli(),

data2 = list(A = A),

iter = 4000,

chains = 4,

cores = 4,

control = list(adapt\_delta = 0.95)

)

ssb\_pred <- fitted(ssb\_fit, summary = TRUE)[, "Estimate"]

df$SSB\_predicted <- ssb\_pred

#Mating System

which(is.na(df$mating\_system))

df[which(is.na(df$mating\_system)), c("nexus\_name", "mating\_system")]

df$mating\_system[df$nexus\_name == "Avahi\_peyrierasi\_ott3607508"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Avahi\_ramanantsoavani\_ott6145314"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Avahi\_unicolor\_ott132960"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Cacajao\_ayresi\_ott6557"] <- "polygynandrous"

df$mating\_system[df$nexus\_name == "Euoticus\_elegantulus\_ott323064"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Euoticus\_pallidus\_ott3607501"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Cheirogaleus\_lavasoensis\_ott5839495"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Phaner\_furcifer\_ott836057"] <- "monogamous"

df$mating\_system[df$nexus\_name == "Pygathrix\_cinerea\_ott786"] <- "polygynous"

df$mating\_system[df$nexus\_name == "Rungwecebus\_kipunji\_ott665522"] <- "polygynandrous"

df$mating\_system[df$nexus\_name == "Pongo\_pygmaeus\_ott770302"] <- "polygynous"

df$mating\_system[df$nexus\_name == "Pongo\_tapanuliensis\_ott7067552"] <- "polygynous"

#Group Structure

which(is.na(df$group\_structure))

df[which(is.na(df$group\_structure)), c("nexus\_name", "group\_structure")]

df$group\_structure[df$nexus\_name == "Avahi\_peyrierasi\_ott3607508"] <- "pair-living"

df$group\_structure[df$nexus\_name == "Avahi\_ramanantsoavani\_ott6145314"] <- "pair-living"

df$group\_structure[df$nexus\_name == "Avahi\_unicolor\_ott132960"] <- "pair-living"

df$group\_structure[df$nexus\_name == "Cacajao\_ayresi\_ott6557"] <- "fission-fusion"

df$group\_structure[df$nexus\_name == "Hapalemur\_occidentalis\_ott311692"] <- "multi-male-multi-female"

#Infant Rearing Strategy

which(is.na(df$infant\_rearing\_strategy))

df[which(is.na(df$infant\_rearing\_strategy)), c("nexus\_name", "infant\_rearing\_strategy")]

df$infant\_rearing\_strategy[df$nexus\_name == "Avahi\_ramanantsoavani\_ott6145314"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Avahi\_unicolor\_ott132960"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Cheirogaleus\_minusculus\_ott311696"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Cheirogaleus\_thomasi\_ott6145312"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Eulemur\_coronatus\_ott394965"] <- "shared-parental care"

df$infant\_rearing\_strategy[df$nexus\_name == "Propithecus\_perrieri\_ott799550"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Phaner\_furcifer\_ott836057"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Rungwecebus\_kipunji\_ott665522"] <- "allo-parental"

df$infant\_rearing\_strategy[df$nexus\_name == "Galago\_gabonensis\_ott323060"] <- "maternal"

df$infant\_rearing\_strategy[df$nexus\_name == "Trachypithecus\_barbei\_ott597818"] <- "allo-parental"

# Hierarchical Structure

which(is.na(df$hierarchical\_structure))

df[which(is.na(df$hierarchical\_structure)), c("nexus\_name", "hierarchical\_structure")]

# === Correctly Fill missing Hierarchical Structure ===

df$hierarchical\_structure[df$nexus\_name == "Avahi\_peyrierasi\_ott3607508"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Avahi\_ramanantsoavani\_ott6145314"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Avahi\_unicolor\_ott132960"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Cacajao\_ayresi\_ott6557"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Cacajao\_hosomi\_ott6559"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Cacajao\_melanocephalus\_ott1022358"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Pygathrix\_cinerea\_ott786"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Cheirogaleus\_sibreei\_ott682744"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Cheirogaleus\_crossleyi\_ott418718"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Phaner\_furcifer\_ott836057"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Galago\_alleni\_ott173071"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Galago\_gabonensis\_ott323060"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Galago\_gallarum\_ott552578"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Galago\_granti\_ott323055"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Hapalemur\_griseus\_ott789630"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Hapalemur\_aureus\_ott671260"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Propithecus\_coquereli\_ott402626"] <- "matriarchal"

df$hierarchical\_structure[df$nexus\_name == "Eulemur\_rufifrons\_ott3607541"] <- "matriarchal"

df$hierarchical\_structure[df$nexus\_name == "Loris\_lydekkerianus\_ott673378"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Loris\_tardigradus\_ott152022"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_pygmaeus\_ott101793"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_javanicus\_ott946857"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_coucang\_ott339043"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_bengalensis\_ott1049687"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_bancanus\_ott6145319"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_borneanus\_ott6145320"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Nycticebus\_kayan\_ott6145321"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_dentatus\_ott712758"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_fuscus\_ott5925595"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_lariang\_ott318683"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_pelengensis\_ott3607736"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_pumilus\_ott3607739"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_sangirensis\_ott49187"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_spectrumgurskyae\_ott7655127"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_supriatnai\_ott7655128"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_tarsier\_ott721280"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_tumpara\_ott6145231"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Tarsius\_wallacei\_ott234367"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Arctocebus\_calabarensis\_ott323051"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Arctocebus\_aureus\_ott673376"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Perodicticus\_potto\_ott339040"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Symphalangus\_syndactylus\_ott417961"] <- "dominant pair"

df$hierarchical\_structure[df$nexus\_name == "Rungwecebus\_kipunji\_ott665522"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Saimiri\_vanzolinii\_ott3607558"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Ateles\_geoffroyi\_ott791691"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Cebus\_kaapori\_ott913292"] <- "patriarchal"

df$hierarchical\_structure[df$nexus\_name == "Cebus\_leucocephalus\_ott6145185"] <- "patriarchal"

df$hierarchical\_structure[df$nexus\_name == "Allenopithecus\_nigroviridis\_ott721471"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Lophocebus\_albigena\_ott554556"] <- "patriarchal"

df$hierarchical\_structure[df$nexus\_name == "Semnopithecus\_schistaceus\_ott3607625"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Semnopithecus\_hector\_ott737534"] <- "egalitarian"

df$hierarchical\_structure[df$nexus\_name == "Papio\_hamadryas\_ott115463"] <- "patriarchal"

df$mating\_system <- as.numeric(as.factor(df$mating\_system))

df$group\_structure <- as.numeric(as.factor(df$group\_structure))

df$infant\_rearing\_strategy <- as.numeric(as.factor(df$infant\_rearing\_strategy))

df$hierarchical\_structure <- as.numeric(as.factor(df$hierarchical\_structure))

df$median\_body\_size <- as.numeric(df$median\_body\_size)

sum(is.na(df$median\_body\_size))

df[is.na(df$median\_body\_size), "nexus\_name"]

# Alouatta species (~7 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Alouatta\_juara\_ott3607545", "Alouatta\_ululata\_ott3607551"

)] <- 7

# Aotus species (~1.2 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Aotus\_brumbacki\_ott338636", "Aotus\_jorgehernandezi\_ott3607598"

)] <- 1.2

# Callicebus and Cheracebus (~1.3 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Callicebus\_barbarabrownae\_ott3607590", "Callicebus\_dubius\_ott3607587",

"Cheracebus\_lugens\_ott927108", "Cheracebus\_medemi\_ott3607588",

"Cheracebus\_purinus\_ott3607591", "Cheracebus\_regulus\_ott3607583"

)] <- 1.3

# Cebus species (~3.5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Cebus\_cesarae\_ott6145182", "Cebus\_malitiosus\_ott6145186",

"Cebus\_unicolor\_ott6145189", "Cebus\_versicolor\_ott6145190",

"Cebus\_yuracus\_ott6145191"

)] <- 3.5

# Cercopithecus species (~5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Cercopithecus\_buettikoferi\_ott7655020", "Cercopithecus\_doggetti\_ott3607605"

)] <- 5

# Cheirogaleus species (~0.5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Cheirogaleus\_andysabini\_ott6145311", "Cheirogaleus\_shethi\_ott7655171"

)] <- 0.5

# Chiropotes species (~3.5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Chiropotes\_israelita\_ott571889", "Chiropotes\_utahickae\_ott571892"

)] <- 3.5

# Chlorocebus (~5 kg)

df$median\_body\_size[df$nexus\_name == "Chlorocebus\_djamdjamensis\_ott875546"] <- 5

# Erythrocebus (~10 kg)

df$median\_body\_size[df$nexus\_name == "Erythrocebus\_poliophaeus\_ott7655031"] <- 10

# Galago and Galagoides (~0.3 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Galago\_minor\_ott7655190", "Galagoides\_orinus\_ott679563",

"Galagoides\_rondoensis\_ott3607491"

)] <- 0.3

# Hoolock (~6 kg)

df$median\_body\_size[df$nexus\_name == "Hoolock\_leuconedys\_ott574262"] <- 6

# Lagothrix (~8 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Lagothrix\_cana\_ott968680", "Lagothrix\_lugens\_ott57134",

"Lagothrix\_poeppigii\_ott57313", "Lagothrix\_tschudii\_ott7067563"

)] <- 8

# Leontopithecus (~0.7 kg)

df$median\_body\_size[df$nexus\_name == "Leontopithecus\_chrysopygus\_ott566865"] <- 0.7

# Lepilemur (~1 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Lepilemur\_aeeclis\_ott757736", "Lepilemur\_ahmansonorum\_ott3607514",

"Lepilemur\_betsileo\_ott3607511", "Lepilemur\_edwardsi\_ott148503",

"Lepilemur\_fleuretae\_ott3607513", "Lepilemur\_grewcockorum\_ott5839498",

"Lepilemur\_hollandorum\_ott6145316", "Lepilemur\_jamesorum\_ott3428",

"Lepilemur\_manasamody\_ott28538", "Lepilemur\_milanoii\_ott3607518",

"Lepilemur\_mitsinjoensis\_ott68775", "Lepilemur\_mittermeieri\_ott441927",

"Lepilemur\_otto\_ott28534", "Lepilemur\_petteri\_ott3607512",

"Lepilemur\_scottorum\_ott6145317", "Lepilemur\_seali\_ott68773",

"Lepilemur\_wrightae\_ott3607515"

)] <- 1

# Macaca speciosa (~6.5 kg)

df$median\_body\_size[df$nexus\_name == "Macaca\_speciosa\_ott115444"] <- 6.5

# Mico species (~0.4 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Mico\_acariensis\_ott3607567", "Mico\_intermedius\_ott3607563",

"Mico\_leucippe\_ott3607564", "Mico\_marcai\_ott3607565",

"Mico\_mauesi\_ott341884", "Mico\_munduruku\_ott7067567",

"Mico\_nigriceps\_ott3607568", "Mico\_rondoni\_ott495388",

"Mico\_saterei\_ott172747"

)] <- 0.4

# Microcebus (~0.05 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Microcebus\_arnholdi\_ott5149", "Microcebus\_bongolavensis\_ott481551",

"Microcebus\_boraha\_ott7655173", "Microcebus\_danfossi\_ott481550",

"Microcebus\_ganzhorni\_ott7067589", "Microcebus\_jollyae\_ott532147",

"Microcebus\_macarthurii\_ott20534", "Microcebus\_mamiratra\_ott560998",

"Microcebus\_manitatra\_ott7655174", "Microcebus\_margotmarshae\_ott5146",

"Microcebus\_marohita\_ott4117712", "Microcebus\_mittermeieri\_ott387144",

"Microcebus\_sambiranensis\_ott765356", "Microcebus\_simmonsi\_ott133625",

"Microcebus\_tanosi\_ott4117710"

)] <- 0.05

# Nycticebus (~1 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Nycticebus\_bancanus\_ott6145319", "Nycticebus\_borneanus\_ott6145320"

)] <- 1

# Otolemur (~1.2 kg)

df$median\_body\_size[df$nexus\_name == "Otolemur\_monteiri\_ott679564"] <- 1.2

# Perodicticus (~1.3 kg)

df$median\_body\_size[df$nexus\_name == "Perodicticus\_potto\_ott339040"] <- 1.3

# Phaner (~0.4 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Phaner\_electromontis\_ott3607530", "Phaner\_parienti\_ott3607529"

)] <- 0.4

# Piliocolobus (~7 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Piliocolobus\_bouvieri\_ott6145121", "Piliocolobus\_lulindicus\_ott7655037",

"Piliocolobus\_oustaleti\_ott6145123", "Piliocolobus\_waldronae\_ott6145127"

)] <- 7

# Pithecia (~2.5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Pithecia\_cazuzai\_ott6145208", "Pithecia\_hirsuta\_ott6145210",

"Pithecia\_inusta\_ott6145211", "Pithecia\_isabela\_ott6145212",

"Pithecia\_milleri\_ott6145213", "Pithecia\_mittermeieri\_ott6145214",

"Pithecia\_pissinattii\_ott6145216", "Pithecia\_rylandsi\_ott6145217",

"Pithecia\_vanzolinii\_ott6145218"

)] <- 2.5

# Plecturocebus (~1.3 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Plecturocebus\_aureipalatii\_ott3607579", "Plecturocebus\_baptista\_ott3607589",

"Plecturocebus\_bernhardi\_ott3607585", "Plecturocebus\_cinerascens\_ott3607584",

"Plecturocebus\_grovesi\_ott7067584", "Plecturocebus\_hoffmannsi\_ott442910",

"Plecturocebus\_miltoni\_ott5925591", "Plecturocebus\_modestus\_ott3607580",

"Plecturocebus\_olallae\_ott3607581", "Plecturocebus\_stephennashi\_ott3607574",

"Plecturocebus\_toppini\_ott7655124", "Plecturocebus\_urubambensis\_ott6145205"

)] <- 1.3

# Presbytis species (~7 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Presbytis\_canicrus\_ott6145130", "Presbytis\_frontata\_ott704371"

)] <- 7

# Pseudopotto (~1 kg)

df$median\_body\_size[df$nexus\_name == "Pseudopotto\_martini\_ott4941439"] <- 1

# Saguinus (~0.6 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Saguinus\_cruzlimai\_ott5852670", "Saguinus\_inustus\_ott765147",

"Saguinus\_leucopus\_ott920962", "Saguinus\_martinsi\_ott156862",

"Saguinus\_melanoleucus\_ott429033", "Saguinus\_nigricollis\_ott1015216",

"Saguinus\_tripartitus\_ott394176", "Saguinus\_ursulus\_ott6145175"

)] <- 0.6

# Sciurocheirus (~0.8 kg)

df$median\_body\_size[df$nexus\_name == "Sciurocheirus\_makandensis\_ott6145318"] <- 0.8

# Tarsius (~0.1 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Tarsius\_dentatus\_ott712758", "Tarsius\_lariang\_ott318683",

"Tarsius\_pelengensis\_ott3607736", "Tarsius\_spectrumgurskyae\_ott7655127",

"Tarsius\_supriatnai\_ott7655128", "Tarsius\_wallacei\_ott234367"

)] <- 0.1

# Trachypithecus (~7.5 kg)

df$median\_body\_size[df$nexus\_name %in% c(

"Trachypithecus\_barbei\_ott597818", "Trachypithecus\_ebenus\_ott3607628",

"Trachypithecus\_shortridgei\_ott514267"

)] <- 7.5

df[which(is.na(df$body\_size\_range)), "nexus\_name"]

# Macaca\_speciosa

df$body\_size\_range[df$nexus\_name == "Macaca\_speciosa\_ott115444"] <- 4.0

# Mico species (~0.8 kg range)

mico\_species <- c(

"Mico\_acariensis\_ott3607567", "Mico\_intermedius\_ott3607563", "Mico\_leucippe\_ott3607564",

"Mico\_marcai\_ott3607565", "Mico\_mauesi\_ott341884", "Mico\_munduruku\_ott7067567",

"Mico\_nigriceps\_ott3607568", "Mico\_rondoni\_ott495388", "Mico\_saterei\_ott172747"

)

df$body\_size\_range[df$nexus\_name %in% mico\_species] <- 0.8

# Microcebus species (~0.5 kg range)

microcebus\_species <- c(

"Microcebus\_arnholdi\_ott5149", "Microcebus\_bongolavensis\_ott481551", "Microcebus\_boraha\_ott7655173",

"Microcebus\_danfossi\_ott481550", "Microcebus\_ganzhorni\_ott7067589", "Microcebus\_jollyae\_ott532147",

"Microcebus\_macarthurii\_ott20534", "Microcebus\_mamiratra\_ott560998", "Microcebus\_manitatra\_ott7655174",

"Microcebus\_margotmarshae\_ott5146", "Microcebus\_marohita\_ott4117712", "Microcebus\_mittermeieri\_ott387144",

"Microcebus\_sambiranensis\_ott765356", "Microcebus\_simmonsi\_ott133625", "Microcebus\_tanosi\_ott4117710"

)

df$body\_size\_range[df$nexus\_name %in% microcebus\_species] <- 0.5

# Nycticebus (~1.0 kg range)

df$body\_size\_range[df$nexus\_name == "Nycticebus\_bancanus\_ott6145319"] <- 1.0

df$body\_size\_range[df$nexus\_name == "Nycticebus\_borneanus\_ott6145320"] <- 1.0

# Otolemur (~1.0 kg)

df$body\_size\_range[df$nexus\_name == "Otolemur\_monteiri\_ott679564"] <- 1.0

# Perodicticus (~1.2 kg)

df$body\_size\_range[df$nexus\_name == "Perodicticus\_potto\_ott339040"] <- 1.2

# Phaner (~0.3 kg)

phaner\_species <- c("Phaner\_electromontis\_ott3607530", "Phaner\_furcifer\_ott836057", "Phaner\_parienti\_ott3607529")

df$body\_size\_range[df$nexus\_name %in% phaner\_species] <- 0.3

# Piliocolobus (~8 kg)

piliocolobus\_species <- c(

"Piliocolobus\_bouvieri\_ott6145121", "Piliocolobus\_lulindicus\_ott7655037",

"Piliocolobus\_oustaleti\_ott6145123", "Piliocolobus\_waldronae\_ott6145127"

)

df$body\_size\_range[df$nexus\_name %in% piliocolobus\_species] <- 8.0

# Pithecia (~2.0 kg)

pithecia\_species <- c(

"Pithecia\_cazuzai\_ott6145208", "Pithecia\_hirsuta\_ott6145210", "Pithecia\_inusta\_ott6145211",

"Pithecia\_isabela\_ott6145212", "Pithecia\_milleri\_ott6145213", "Pithecia\_mittermeieri\_ott6145214",

"Pithecia\_pissinattii\_ott6145216", "Pithecia\_rylandsi\_ott6145217", "Pithecia\_vanzolinii\_ott6145218"

)

df$body\_size\_range[df$nexus\_name %in% pithecia\_species] <- 2.0

# Plecturocebus (~1.5 kg)

plecturocebus\_species <- c(

"Plecturocebus\_aureipalatii\_ott3607579", "Plecturocebus\_baptista\_ott3607589", "Plecturocebus\_bernhardi\_ott3607585",

"Plecturocebus\_cinerascens\_ott3607584", "Plecturocebus\_cupreus\_ott34118", "Plecturocebus\_grovesi\_ott7067584",

"Plecturocebus\_hoffmannsi\_ott442910", "Plecturocebus\_miltoni\_ott5925591", "Plecturocebus\_modestus\_ott3607580",

"Plecturocebus\_olallae\_ott3607581", "Plecturocebus\_stephennashi\_ott3607574", "Plecturocebus\_toppini\_ott7655124",

"Plecturocebus\_urubambensis\_ott6145205"

)

df$body\_size\_range[df$nexus\_name %in% plecturocebus\_species] <- 1.5

# Presbytis (~5 kg)

presbytis\_species <- c(

"Presbytis\_canicrus\_ott6145130", "Presbytis\_frontata\_ott704371",

"Presbytis\_pileata\_ott7655041", "Presbytis\_potenziani\_ott704401"

)

df$body\_size\_range[df$nexus\_name %in% presbytis\_species] <- 5.0

# Pseudopotto (~1.5 kg)

df$body\_size\_range[df$nexus\_name == "Pseudopotto\_martini\_ott4941439"] <- 1.5

# Saguinus (~0.7 kg)

saguinus\_species <- c(

"Saguinus\_cruzlimai\_ott5852670", "Saguinus\_graellsi\_ott955071", "Saguinus\_inustus\_ott765147",

"Saguinus\_leucopus\_ott920962", "Saguinus\_martinsi\_ott156862", "Saguinus\_melanoleucus\_ott429033",

"Saguinus\_nigricollis\_ott1015216", "Saguinus\_tripartitus\_ott394176", "Saguinus\_ursulus\_ott6145175"

)

df$body\_size\_range[df$nexus\_name %in% saguinus\_species] <- 0.7

# Sciurocheirus (~1.0 kg)

df$body\_size\_range[df$nexus\_name == "Sciurocheirus\_makandensis\_ott6145318"] <- 1.0

# Tarsius (~0.2 kg)

tarsius\_species <- c(

"Tarsius\_dentatus\_ott712758", "Tarsius\_lariang\_ott318683", "Tarsius\_pelengensis\_ott3607736",

"Tarsius\_spectrumgurskyae\_ott7655127", "Tarsius\_supriatnai\_ott7655128", "Tarsius\_wallacei\_ott234367"

)

df$body\_size\_range[df$nexus\_name %in% tarsius\_species] <- 0.2

# Trachypithecus (~5 kg)

trachypithecus\_species <- c(

"Trachypithecus\_barbei\_ott597818", "Trachypithecus\_ebenus\_ott3607628", "Trachypithecus\_shortridgei\_ott514267"

)

df$body\_size\_range[df$nexus\_name %in% trachypithecus\_species] <- 5.0

predicted\_data <- df %>%

select(

nexus\_name,

SSB\_predicted,

ends\_with("\_pred"),

mating\_system,

group\_structure,

infant\_rearing\_strategy,

hierarchical\_structure

)

write.csv(predicted\_data, "Bayesian\_trait\_predictions.csv", row.names = FALSE)

df\_scaled <- df

numeric\_vars <- sapply(df\_scaled, is.numeric)

numeric\_vars["SSB\_binary"] <- FALSE

df\_scaled[numeric\_vars] <- scale(df\_scaled[numeric\_vars])

model\_hierarchical <- '

# SSB presence

SSB\_latent =~ SSB\_predicted

# Social predictors of SSB

SSB\_latent ~ group\_structure + mating\_system + infant\_rearing\_strategy + hierarchical\_structure + median\_group\_size

# Life history predictors of social structure

group\_structure ~ lifespan\_resid\_pred + sdimorphism\_resid\_pred + median\_body\_size + body\_size\_range

mating\_system ~ lifespan\_resid\_pred + sdimorphism\_resid\_pred + median\_body\_size + body\_size\_range

infant\_rearing\_strategy ~ lifespan\_resid\_pred + sdimorphism\_resid\_pred + median\_body\_size + body\_size\_range

hierarchical\_structure ~ lifespan\_resid\_pred + sdimorphism\_resid\_pred + median\_body\_size + body\_size\_range

median\_group\_size ~ lifespan\_resid\_pred + sdimorphism\_resid\_pred + median\_body\_size + body\_size\_range

# Environmental predictors of life history

lifespan\_resid\_pred ~ temp\_resid\_pred + precip\_resid\_pred + cloud\_resid\_pred + biomass\_resid\_pred + predation\_resid\_pred

sdimorphism\_resid\_pred ~ temp\_resid\_pred + precip\_resid\_pred + cloud\_resid\_pred + biomass\_resid\_pred + predation\_resid\_pred

median\_body\_size ~ temp\_resid\_pred + precip\_resid\_pred + cloud\_resid\_pred + biomass\_resid\_pred + predation\_resid\_pred

body\_size\_range ~ temp\_resid\_pred + precip\_resid\_pred + cloud\_resid\_pred + biomass\_resid\_pred + predation\_resid\_pred

'

fit\_path\_analysis <- sem(

model\_hierarchical,

data = df\_scaled,

estimator = "ML",

missing = "fiml"

)

summary(fit\_path\_analysis, standardized = TRUE, fit.measures = TRUE)

######################################

##Species Similarity Checks

install.packages("dplyr")

install.packages("vegan")

install.packages("ggplot2")

library(dplyr)

library(vegan)

library(ggplot2)

library(tidyr)

df <- read.csv("Species list comparison.csv")

Species\_Search1 <- df$sexual\_behaviour\_search

Mentions\_1 <- df$sexual\_behavious\_mentions

Species\_Search2 <- df$ssb\_search

Mentions\_2 <- df$ssb\_mentions

#Jaccard Similarity Index

df <- df %>%

mutate(Presence\_1 = ifelse(is.na(Mentions\_1), 0, ifelse(Mentions\_1 > 0, 1, 0)),

Presence\_2 = ifelse(is.na(Mentions\_2), 0, ifelse(Mentions\_2 > 0, 1, 0)))

intersection <- sum(df$Presence\_1 & df$Presence\_2)

union <- sum(df$Presence\_1 | df$Presence\_2)

jaccard\_similarity <- intersection / union

print(paste("Jaccard similarity index:", round(jaccard\_similarity, 3)))

#Fisher Test

contingency\_table <- table(df$Presence\_1, df$Presence\_2)

fisher\_test <- fisher.test(contingency\_table)

print(fisher\_test)

#Spearman's rank correlation

spearman\_corr <- cor.test(Mentions\_1, Mentions\_2, method = "spearman")

print(spearman\_corr)