

```
setwd("~/Documents/Working Directory/PNAS")
rm(list=ls())
```

```
#####
```

```
# Libraries #
```

```
#####
```

```
library("tidyverse")
```

```
library("rjags")
```

```
library("coda")
```

```
library("matrixStats")
```

```
library("ggplot2")
```

```
#####
```

```
# #
```

```
# 3. PREPARE EC REARING DATA #
```

```
# #
```

```
#####
```

```
#####
```

```
# Read in Data #
```

```
#####
```

```
#load rearing data
```

```

df<- read.csv("EC.csv", header=TRUE, stringsAsFactors = FALSE) ##rearing data

#####

# Fix dates

#####

df$date <- as.Date(df$date, format = "%d-%b-%y")

#Add year, month , day columns to allow for various time analyses

df$year <- as.numeric(format(df$date, "%Y")) # year
df$month <- as.numeric(format(df$date, "%m")) # month
df$day <- as.numeric(format(df$date, "%d")) # day

#Remove data for which there is no date
df<-df[complete.cases(df$date),]
df<- df[order(df$date),]

#Remove coleopteran and non-parasitoid hymenopteran species from database
df <- df[ !(df$insectorder %in% c("Coleoptera","Hymenoptera")), ]
df<-droplevels(df)

#####

# Plot ("plot" or "gencol") #####

#####

#Designate factor called "collect_method" which has two levels "plot" or "gencol"
#Plot means data was collected from standardized plot
#gencol means data was collected via haphazard collection
df$collect_method<- ifelse(grepl("plot", df$loc, ignore.case = F), "plot", "gencol")
df$collect_method[grepl("Plot",df$loc)]<-"plot"

#Add category that specifies lep genera

```

```
df$lepgen<-gsub( ".*$", "", df$lepsp )
```

```
#subset and read out df
```

```
file<- df[,c("ID", "loc", "date","day", "month", "year", "lepsp", "lepfam", "collect_method", "lepgen")]
```

```
file$lepsp<-as.factor(file$lepsp)
```

```
file$lepfam<-as.factor(file$lepfam)
```

```
write.csv(file,"EC_workingfile.csv")
```

```
#####
```

```
# file: Add super family column
```

```
#####
```

```
#remove non-yanayacu locations
```

```
file<-file[!grepl("ahuano", file$loc),]
```

```
file<-file[!grepl("km 33 via al tena", file$loc),]
```

```
file<-file[!grepl("baeza", file$loc),]
```

```
file<-file[!grepl("richmanyacu", file$loc),]
```

```
file<-file[!grepl("baeza road", file$loc),]
```

```
file<-file[!grepl("tena", file$loc),]
```

```
file<-file[!grepl("camino a loreto", file$loc),]
```

```
file<-file[!grepl("chaco", file$loc),]
```

```
file<-file[!grepl("papallacta", file$loc),]
```

```
file<-file[!grepl("cayambe coca", file$loc),]
```

```
file<-file[!grepl("rio chalpi grande", file$loc),]
```

```
file<-file[!grepl("rio arenillas", file$loc),]
```

```
file<-file[!grepl("alinahui road", file$loc),]
```

```
file<-file[!grepl("chontaloma road", file$loc),]
```

```
file<-file[!grepl("rio chontayacu", file$loc),]
```

file<-file[!grep("narupa km 24", file\$loc),]  
file<-file[!grep("laguna papallacta", file\$loc),]  
file<-file[!grep("camino de las termas de papallacta", file\$loc),]  
file<-file[!grep("shiripuno", file\$loc),]  
file<-file[!grep("cocodrilos", file\$loc),]  
file<-file[!grep("antenas del chaco", file\$loc),]  
file<-file[!grep("baeza granja integral", file\$loc),]  
file<-file[!grep("rio oyacachi el chaco", file\$loc),]  
file<-file[!grep("sendero antenas tena", file\$loc),]  
file<-file[!grep("jatun sacha", file\$loc),]  
file<-file[!grep("rio san juan el chaco", file\$loc),]  
file<-file[!grep("wasca yacu via al tena", file\$loc),]  
file<-file[!grep("rio malo", file\$loc),]  
file<-file[!grep("ahuano", file\$loc),]  
file<-file[!grep("quijos huaico", file\$loc),]  
file<-file[!grep("yasuni", file\$loc),]  
file<-file[!grep("narupa sendero el jibaro", file\$loc),]  
file<-file[!grep("rio palenque reserve sendero 1", file\$loc),]  
file<-file[!grep("rio palenque reserve sendero 3", file\$loc),]  
file<-file[!grep("rio anzu reserve mera ecuador", file\$loc),]  
file<-file[!grep("pungarayacu", file\$loc),]  
file<-file[!grep("antenas baeza", file\$loc),]  
file<-file[!grep("iyrania rio napo", file\$loc),]  
file<-file[!grep("Reserva Rio Zunac", file\$loc),]  
file<-file[!grep("Baeza-Cascadas", file\$loc),]  
file<-file[!grep("Reserva las Candelaria", file\$loc),]  
file<-file[!grep("Rio Zunac", file\$loc),]  
file<-file[!grep("Shiripuno lodge", file\$loc),]  
file<-file[!grep("Iyarina", file\$loc),]

```
file<-file[!grep("Huaroni comunidad", file$loc),]
```

```
file<-file[!grep("Narupayacu", file$loc),]
```

```
file<-file[!grep("plot15", file$loc),]
```

```
file<-file[!grep("plot16", file$loc),]
```

```
file<-file[!grep("plot20", file$loc),]
```

```
file<-file[!grep("plot23", file$loc),]
```

```
file<-file[!grep("plot24", file$loc),]
```

```
file<-file[!grep("plot37", file$loc),]
```

```
file<-file[!grep("plot38", file$loc),]
```

```
file<-file[!grep("plot39", file$loc),]
```

```
file<-file[!grep("plot40", file$loc),]
```

```
file<-file[!grep("plot42", file$loc),]
```

```
file<-file[!grep("plot43", file$loc),]
```

```
file<-file[!grep("plot44", file$loc),]
```

```
file<-file[!grep("plot45", file$loc),]
```

```
file<-file[!grep("plot46", file$loc),]
```

```
file<-file[!grep("plot51", file$loc),]
```

```
file<-file[!grep("plot52", file$loc),]
```

```
file<-file[!grep("plot53", file$loc),]
```

```
file<-file[!grep("plot54", file$loc),]
```

```
file<-file[!grep("plot55", file$loc),]
```

```
file<-file[!grep("plot56", file$loc),]
```

```
file<-file[!grep("plot57", file$loc),]
```

```
file<-file[!grep("plot58", file$loc),]
```

```
file<-file[!grep("plot59", file$loc),]
```

```
file<-file[!grep("plot60", file$loc),]
```

```
file<-file[!grep("plot61", file$loc),]
```

```
file<-file[!grep("plot62", file$loc),]
```

file<-file[!grep("plot63", file\$loc),]  
file<-file[!grep("plot64", file\$loc),]  
file<-file[!grep("plot66", file\$loc),]  
file<-file[!grep("plot67", file\$loc),]  
file<-file[!grep("plot68", file\$loc),]  
file<-file[!grep("plot69", file\$loc),]  
file<-file[!grep("plot71", file\$loc),]  
file<-file[!grep("plot84", file\$loc),]  
file<-file[!grep("plot85", file\$loc),]  
file<-file[!grep("plot92", file\$loc),]  
file<-file[!grep("plot107", file\$loc),]  
file<-file[!grep("plot113", file\$loc),]  
file<-file[!grep("plot115", file\$loc),]  
file<-file[!grep("plot126", file\$loc),]  
file<-file[!grep("plot127", file\$loc),]  
file<-file[!grep("plot134", file\$loc),]  
file<-file[!grep("plot135", file\$loc),]  
file<-file[!grep("plot139", file\$loc),]  
file<-file[!grep("plot167", file\$loc),]  
file<-file[!grep("plot168", file\$loc),]  
file<-file[!grep("plot190", file\$loc),]  
file<-file[!grep("plot196", file\$loc),]  
file<-file[!grep("plot197", file\$loc),]  
file<-file[!grep("plot198", file\$loc),]  
file<-file[!grep("plot199", file\$loc),]  
file<-file[!grep("plot202", file\$loc),]  
file<-file[!grep("plot210", file\$loc),]  
file<-file[!grep("plot215", file\$loc),]  
file<-file[!grep("plot221", file\$loc),]

```
file<-file[!grep("plot222", file$loc),]
file<-file[!grep("plot237", file$loc),]
file<-file[!grep("plot238", file$loc),]
file<-file[!grep("plot247", file$loc),]
file<-file[!grep("plot248", file$loc),]
file<-file[!grep("plot253", file$loc),]
file<-file[!grep("plot254", file$loc),]
file<-file[!grep("plot257", file$loc),]
file<-file[!grep("plot260", file$loc),]
file<-file[!grep("plot261", file$loc),]
file<-file[!grep("plot279", file$loc),]
file<-file[!grep("plot280", file$loc),]
file<-file[!grep("plot288", file$loc),]
file<-file[!grep("plot304", file$loc),]
file<-file[!grep("plot305", file$loc),]
file<-file[!grep("plot306", file$loc),]
file<-file[!grep("plot308", file$loc),]
file<-file[!grep("plot309", file$loc),]
file<-file[!grep("plot319", file$loc),]
file<-file[!grep("plot320", file$loc),]
file<-file[!grep("plot315", file$loc),]
file<-file[!grep("plot316", file$loc),]
file<-file[!grep("plot322", file$loc),]
file<-file[!grep("plot323", file$loc),]
file<-file[!grep("plot324", file$loc),]
file<-file[!grep("plot325", file$loc),]
file<-file[!grep("plot326", file$loc),]
file<-file[!grep("plot327", file$loc),]
file<-file[!grep("plot330", file$loc),]
```

```
file<-file[!grep("plot331", file$loc),]
file<-file[!grep("plot332", file$loc),]
file<-file[!grep("plot346", file$loc),]
file<-file[!grep("plot393", file$loc),]
file<-file[!grep("plot44", file$loc),]
file<-file[!grep("plot527", file$loc),]
file<-file[!grep("plot530", file$loc),]
file<-file[!grep("plot531", file$loc),]
file<-file[!grep("plot554", file$loc),]
file<-file[!grep("plot568", file$loc),]
file<-file[!grep("plot570", file$loc),]
```

#Remove rows with Eois or non-lepidopteran entries

```
file<-file[!grep("eois", file$lepsz),]
file<-file[!grep("Eois", file$lepsz),]
```

```
file<-file[!grep("coleccionenlanoche", file$lepfam),]
file<-file[!grep("tenthredinidae", file$lepfam),]
file<-file[!grep("trichoptera", file$lepfam),]
file<-file[!grep("Trichoptera", file$lepfam),]
file<-file[!grep("chrysomelidae", file$lepfam),]
file<-file[!grep("unknown", file$lepfam),]
file$superfam[file$lepfam=="choreutidae"]<-"Choreutoidea"
file$superfam[file$lepfam=="erebidae" ]<-"Noctuoidea"
file$superfam[file$lepfam=="lasiocampidae" ]<-"Lasiocampoidea"
file$superfam[file$lepfam=="mimallonidae" ]<-"Mimallonoidea"
file$superfam[file$lepfam=="nymphalidae" ]<-"Papilionoidea"
```



file\$superfam[file\$lepfam=="pieridae" ]<-"Papilionoidea"  
file\$superfam[file\$lepfam=="riodinidae"]<-"Papilionoidea"  
file\$superfam[file\$lepfam=="sphingidae"]<-"Bombycoidea"  
file\$superfam[file\$lepfam=="uraniidae"]<-"Geometroidea"  
file\$superfam[file\$lepfam=="aididae"]<-"Zygaenoidea"  
file\$superfam[file\$lepfam=="crambidae"]<-"Pyraloidea"  
file\$superfam[file\$lepfam=="gelechiidae"]<-"Gelechioidea"  
file\$superfam[file\$lepfam=="limacodidae"]<-"Zygaenoidea"  
file\$superfam[file\$lepfam=="noctuidae"]<-"Noctuoidea"  
file\$superfam[file\$lepfam=="oecophoridae"]<-"Gelechioidea"  
file\$superfam[file\$lepfam=="psychidae"]<-"Tineoidea"  
file\$superfam[file\$lepfam=="saturniidae"]<-"Bombycoidea"  
file\$superfam[file\$lepfam=="thyrididae"]<-"Thyridoidea"  
file\$superfam[file\$lepfam=="zygaenidae"]<-"Zygaenoidea"  
file\$superfam[file\$lepfam=="apatelodidae"]<-"Bombycoidea"  
file\$superfam[file\$lepfam=="dalceridae"]<-"Zygaenoidea"  
file\$superfam[file\$lepfam=="geometridae"]<-"Geometroidea"  
file\$superfam[file\$lepfam=="lycaenidae"]<-"Papilionoidea"  
file\$superfam[file\$lepfam=="nolidae"]<-"Noctuoidea"  
file\$superfam[file\$lepfam=="pantheidae"]<-"Noctuoidea"  
file\$superfam[file\$lepfam=="pterophoridae"]<-"Pterophoroidea"  
file\$superfam[file\$lepfam=="scythruidae"]<-"Gelechioidea"  
file\$superfam[file\$lepfam=="tineidae"]<-"Tineoidea"  
file\$superfam[file\$lepfam=="bombycidae"]<-"Bombycoidea"  
file\$superfam[file\$lepfam=="drepanidae"]<-"Drepanoidea"  
file\$superfam[file\$lepfam=="hesperiidae"]<-"Papilionoidea"  
file\$superfam[file\$lepfam=="megalopygidae"]<-"Zygaenoidea"  
file\$superfam[file\$lepfam=="notodontidae"]<-"Noctuoidea"  
file\$superfam[file\$lepfam=="papilionidae"]<-"Papilionoidea"

```

file$superfam[file$lepfam=="pyralidae"]<-"Pyraloidea"
file$superfam[file$lepfam=="sematuridae"]<-"Geometroidea"
file$superfam[file$lepfam=="tortricidae"]<-"Tortricoidea"
file<-droplevels(file)
file$superfam<-as.factor(file$superfam)

```

#----Get an idea of how many samples per super family so that we can create minimum threshold to be included in study

```

temp <-
  file%>%
  filter(!is.na(superfam))%>%
  group_by(year, superfam) %>%
  dplyr::summarize(freq = n())

ggplot(temp, aes(x=year, y=freq, colour=superfam)) +
  geom_point() +
  geom_smooth(method=lm, se=FALSE)+
  facet_wrap(~superfam)

```

#Remove (removed below) the following family bc these superfamilies were samped only 1-3 years:

#Choreutoidea, Gracillariodea, Tineoidea, Pterophoroidea, Tineaidea

```

file<- file[!(file$superfam %in% c("Mimallonoidea", "Thyridoidea", "Tineoidea", "Gelechioidea")),]
file<-droplevels(file)

year<-file
month_df<-file

```

```
#####
```

```
# Summarise frequency table
```

```
#####
```

```
#----ANNUAL SUMMARY
```

```
a_year <-
```

```
  year%>%
```

```
  filter(!is.na(superfam))%>%
```

```
  group_by(year, superfam) %>%
```

```
  dplyr::summarize(freq = n())
```

```
p_year <-
```

```
  year%>%
```

```
  filter(collect_method=="plot", !is.na(superfam))%>%
```

```
  group_by(year, superfam) %>%
```

```
  dplyr::summarize(freq = n())
```

```
g_year <-
```

```
  year%>%
```

```
  filter(collect_method=="gencol", !is.na(superfam))%>%
```

```
  group_by(year, superfam) %>%
```

```
  dplyr::summarize(freq = n())
```

```
#----MONTHLY SUMMARY
```

```
#----Count of super fam freq per month and collection method
```

```
month_df <-
```

```
month_df%>%
filter(!is.na(superfam))%>%
group_by(year, month, collect_method) %>%
dplyr::summarize(freq = n())
```

```
all_methods <-
month_df%>%
group_by(year,month) %>%
dplyr::summarize(freq = sum(freq))
all_methods$collect_method<- "all"
month_df<-rbind(as.data.frame(month_df), as.data.frame(all_methods))
```

```
#####
#                                     #
#           4. BAYESIAN ANALYSES           #
#                                     #
#####
```

```
dat<-a_year
temp<-dat
dat$freq <- as.numeric(scale(log(dat$freq)))
dat$year <- as.numeric(scale(dat$year))
dat<-dat[ which( ! dat$year %in% NA) , ]
```

```
library("rjags")
library("coda")
library("matrixStats")
## put into list format jags wants
```

```
dataForJags <- list(freq=dat$freq, year=dat$year, superfam=dat$superfam,  
Ncol=length(levels(dat$superfam)), N=length(dat$freq))
```

```
## specify model, priors etc.
```

```
model <- "model{
```

```
for(i in 1:N){
```

```
freq[i] ~ dnorm(alpha[i], tau)
```

```
alpha[i] <- mu[superfam[i]] + beta1[superfam[i]]*year[i]
```

```
}
```

```
# coefficients for individual collection_method (conditional priors)
```

```
for(j in 1:Ncol){
```

```
beta1[j] ~ dnorm(beta1mu, beta1tau)
```

```
mu[j] ~ dnorm(mumu, mutau)
```

```
}
```

```
tau ~ dgamma(0.1,0.1)
```

```
# hyperpriors for coefficients
```

```
beta1mu ~ dnorm(0,0.001)
```

```
mumu ~ dnorm(0,0.001)
```

```
# hyperpriors for precisions
```

```
beta1tau ~ dgamma(0.1,0.1)
```

```
mutau ~ dgamma(0.1,0.1)
```

```
}"
```

```
## compile
```

```
mod1 <- jags.model(textConnection(model),data= dataForJags,n.chains=2)
```

```

## run it

samples<-jags.samples(model= mod1,variable.names=c("beta1","mu","beta1mu","mumu",
"beta1tau","mutau","tau"),n.iter=1000000)

#take beta1mu (coefficient and 95%CI for highest level of heirarchy) to later ass to caterpillar plot
keyBits <-round(quantile(samples$beta1mu[,,],probs=c(0.5,0.1,0.9)),3)

#Calculate probability of Negative slope
probNeg<-as.vector(samples$beta1mu[,,1])

probNeg <- length(probNeg[probNeg<0]) / length(as.vector(samples$beta1mu[,,1]))

vec <- data.frame(site= "EC", superfam= "Overall Beta Coefficient", mid=keyBits[1], lower=keyBits[2],
upper=keyBits[3], probNeg=probNeg)

vec

#### generate a table of genus-specific slopes and intervals

mid <- NA

lower <- NA

upper <- NA

probNeg <- NA

res <- samples$beta1

for(i in 1:length(levels(dat$superfam))){
  sub <- res[i,,]

  mid[i] <- quantile(sub,0.5)

  upper[i] <- quantile(sub,0.9)

  lower[i] <- quantile(sub,0.1)

  probNeg[i] <- length(sub[sub<0]) / length(sub)

  print(i)
}

new <- data.frame(levels(dat$superfam),mid,lower,upper, probNeg)

new <- new[order(new$mid,decreasing=T),]

```

```

colnames(new)[1] <- "superfam"
new$site<- "EC"
new<-rbind(new, vec)
#hist(new$mid, col="green",xlim=c(-0.2, 0.0001))
head(new)

write.csv(new, "new.csv")

#####
#                                     #
#           4. GRAPHS                 #
#                                     #
#####

#####
#           Caterpillar Plot
#
#####

#Use for a_year
new$border_color <- c(rep("#51B700",3), rep("#3B94ED", 8),"#1B1E23")
new$fill_color <- c(rep("#51B700",3), rep("#3B94ED", 8),"#1B1E23")

# Use for p_year
new$border_color <- c( rep("#3B16A1", 2), "#3B94ED", "#3B16A1", "#3B94ED", rep("#3B16A1", 6),
"#1B1E23")
new$fill_color <- c( rep("#3B16A1", 2), "#3B94ED", "#3B16A1", "#3B94ED", rep("#3B16A1", 6),
"#1B1E23")

#Use for g_year

```

```

new$border_color <- c(rep("#51B700", 5), rep("#3B94ED", 6), "#1B1E23")
new$fill_color <- c(rep("#51B700", 5), rep("#3B94ED", 6), "#1B1E23")

#use for g_year
quartz(width=7,height=9)

## plot those point estimates and credible intervals in caterpillar plot
layout(matrix(c(1,1,1,2,1,1,1,3,1,1,1,4,1,1,1,5,1,1,1,6),5,4,byrow=T))
par(mar=c(5,1,0,5))
par(oma=c(12,0,1,4))

plot(new$mid,12:1,yaxt="n",ylab="",xlab="",xlim=c(-0.5, 1),ylim=c(1,12),cex.axis=2, cex=3, font=2,
cex.lab=5, pch=21,col=new$border_color, bg= alpha(new$fill_color, 0.5), lwd=2)

mtext("Change in Annual Frequency",1,line=2.5,cex=1.5, font= 2)

arrows(new$lower, 12:1, new$upper, 12:1,length=0,lwd=2, col=new$fill_color)

abline(v=0,lty=2)

axis(4,labels= new$superfam,at=12:1,las=2,cex.axis=2,font=2,tck=-0.02)

grid(nx=10,ny=20)

#####

#   Box Plot: Monthly Freq Across Years

#

#####

# Box Plot: Monthly Freq Across Years

# purple: plot

# blue: gencol

# forest green= all

ec <-

```



```

month_df%>%
group_by(year, collect_method) %>%
dplyr::summarize(mean=mean(freq), sd=sd(freq), n=n_distinct(month))
ec <-
ec%>%
group_by(year, collect_method ) %>%
dplyr::mutate( se=(sd/sqrt(n)))
ec$mean<-scale(ec$mean)
ec$se<-scale(ec$se)
ec$se[is.na(ec$se)] <- 0

ec$year<-as.factor(ec$year)
lvls<-levels(ec$year)

ggplot(ec, aes(x=as.factor(year), y=mean, color= collect_method))+
  geom_pointrange(aes(ymin = mean-se, ymax = mean+se, color=
collect_method),position=position_jitterdodge(-2),alpha=0.8, size=0.75)+
  scale_color_manual(values = c( "#000000", "#CDCDCD", "#878787"))+
  scale_x_discrete(limits=lvls,breaks=lvls[seq(1,length(lvls),by=2)))+
  labs(x= "Year", y=expression(paste("Frequency of Encounter")))+
  theme_classic()+
  theme(axis.text.x = element_text(face= "bold", size= 20, angle=18, hjust = 1),
        axis.text.y =element_text(face= "bold", size=20),
        axis.title.x= element_text( size = rel(2.5)),
        axis.title.y= element_text( size = rel(2.5)))

```