Acroporid gap analysis

2024-10-05

## Exploratory Data Analysis

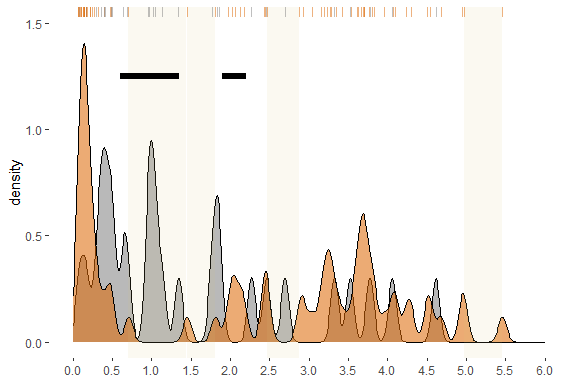
# Load data  
raw\_ages <- read\_excel("raw\_ages.xlsx")  
raw\_mx <- raw\_ages |> filter(source == "This paper")  
  
# Select appropriate bandwidth value  
bw\_rawmx\_ucv <- bw.ucv(raw\_mx$t\_ka)

#> Warning in bw.ucv(raw\_mx$t\_ka): minimum occurred at one end of the range

bw\_rawmx\_ucv

#> [1] 0.07870325

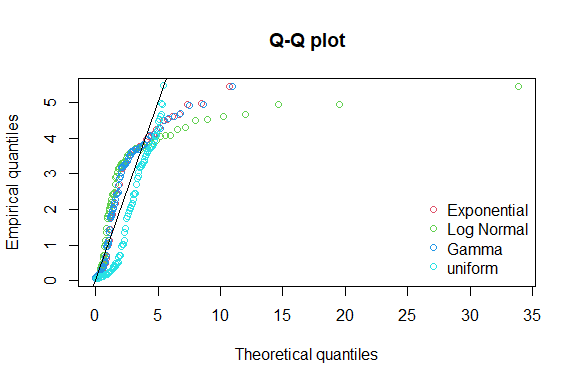
# Plot KDE  
ggplot(raw\_mx) +  
 aes(x = t\_ka) +  
 geom\_density(aes(y = after\_stat(density), fill = group), alpha = 0.55, adjust = bw\_rawmx\_ucv) +  
 scale\_fill\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), legend.position = "none") +  
 scale\_x\_continuous(limits = c(0, 6), breaks = seq(0, 6, 0.5)) +  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 0.5)) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), legend.position = "none", axis.title.x = element\_blank()) +  
 annotate("rect", xmin = 0.71, xmax = 1.44, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#DAC778") +  
 annotate("rect", xmin = 1.46, xmax = 1.81, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#DAC778") +  
 annotate("rect", xmin = 2.47, xmax = 2.88, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#DAC778") +  
 annotate("rect", xmin = 4.97, xmax = 5.46, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#DAC778") +  
 annotate("segment", x = 0.6, xend = 1.35, y = 1.25, yend = 1.25, color = "black", linewidth = 2.25) +  
 annotate("segment", x = 1.9, xend = 2.2, y = 1.25, yend = 1.25, color = "black", linewidth = 2.25) +  
 geom\_rug(aes(y = 1.5, colour = group), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray"))



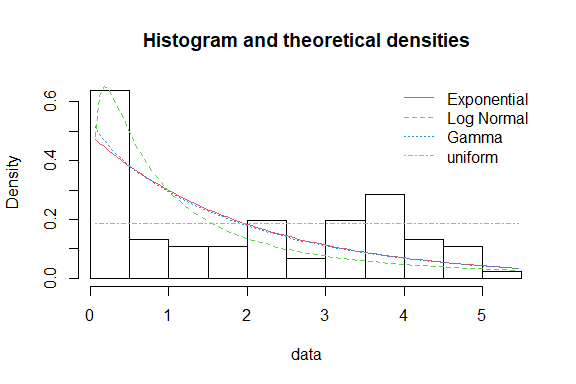
### Null Model selection (Monte Carlo simulation)

#### scenario (1) the entire range of ages

# Exploring theoretical distributions for Monte Carlo simulation  
age <- raw\_ages |> filter(source == "This paper")  
ages <- age$t\_ka  
  
# Fitting multiple distributions  
fit\_exponential <- fitdist(ages, "exp")  
fit\_lognormal <- fitdist(ages, "lnorm")  
fit\_gamma <- fitdist(ages, "gamma")  
fit\_uniforme <- fitdist(ages, "unif")  
  
# Q-Q plot  
par(mfrow = c(1, 1))  
qqcomp(  
 list(fit\_exponential, fit\_lognormal, fit\_gamma, fit\_uniforme),  
 legendtext = c("Exponential", "Log Normal", "Gamma", "uniform")  
)



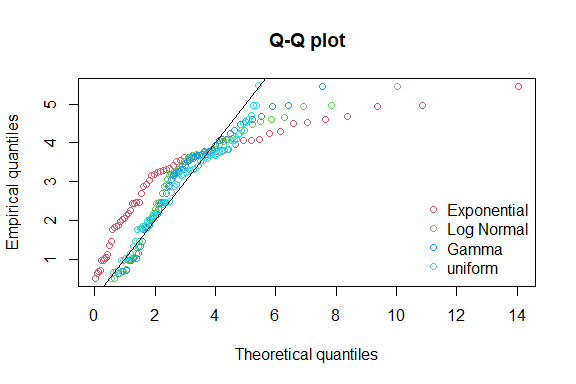
# Histogram with fitted density curves  
denscomp(  
 list(fit\_exponential, fit\_lognormal, fit\_gamma, fit\_uniforme),  
 legendtext = c("Exponential", "Log Normal", "Gamma", "uniform")  
)



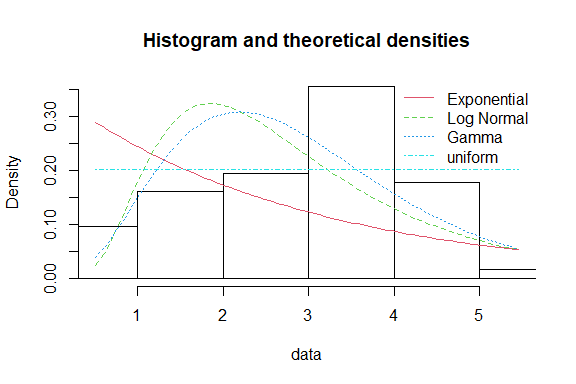
# Model selection with AIC and BIC  
aic\_values <- c(fit\_exponential$aic, fit\_lognormal$aic, fit\_gamma$aic, fit\_uniforme$aic)  
bic\_values <- c(fit\_exponential$bic, fit\_lognormal$bic, fit\_gamma$bic, fit\_uniforme$bic)  
  
Model.full <- data.frame(  
 Data = "All data",  
 Distribucion = c("Exponential", "Log Normal", "Gamma", "Uniform"),  
 AIC = aic\_values,  
 BIC = bic\_values  
)

#### scenario (2) Exclusion of the most recent 500 years

# Exclusion of the most recent 500 years  
ages500 <- age$t\_ka[age$t\_ka > 0.5]  
  
# Fitting multiple distributions  
fit500\_exponential <- fitdist(ages500, "exp")  
fit500\_lognormal <- fitdist(ages500, "lnorm")  
fit500\_gamma <- fitdist(ages500, "gamma")  
fit500\_uniforme <- fitdist(ages500, "unif")  
  
# Q-Q plot  
qqcomp(  
 list(fit500\_exponential, fit500\_lognormal, fit500\_gamma, fit500\_uniforme),  
 legendtext = c("Exponential", "Log Normal", "Gamma", "uniform")  
)



# Histogram with fitted density curves  
denscomp(  
 list(fit500\_exponential, fit500\_lognormal, fit500\_gamma, fit500\_uniforme),  
 legendtext = c("Exponential", "Log Normal", "Gamma", "uniform")  
)



# Model selection with AIC and BIC  
aic\_values500 <- c(fit500\_exponential$aic, fit500\_lognormal$aic, fit500\_gamma$aic, fit500\_uniforme$aic)  
bic\_values500 <- c(fit500\_exponential$bic, fit500\_lognormal$bic, fit500\_gamma$bic, fit500\_uniforme$bic)  
  
Model.500 <- data.frame(  
 Data = "Exclusion of 500 most recent years",  
 Distribucion = c("Exponential", "Log Normal", "Gamma", "Uniform"),  
 AIC = aic\_values500,  
 BIC = bic\_values500  
)

### Null Model analysis results (Best Fit Selection)

models <- bind\_rows(Model.full, Model.500)  
  
ft <- flextable(models) %>%  
 theme\_vanilla() %>%  
 autofit() %>%  
 set\_caption(caption = "Analysis results")  
  
print(ft)

#> a flextable object.  
#> col\_keys: `Data`, `Distribucion`, `AIC`, `BIC`   
#> header has 1 row(s)   
#> body has 8 row(s)   
#> original dataset sample:   
#> Data Distribucion AIC BIC  
#> 1 All data Exponential 315.3666 317.8774  
#> 2 All data Log Normal 336.1135 341.1352  
#> 3 All data Gamma 317.2263 322.2480  
#> 4 All data Uniform 310.7452 315.7669  
#> 5 Exclusion of 500 most recent years Exponential 258.6327 260.7599

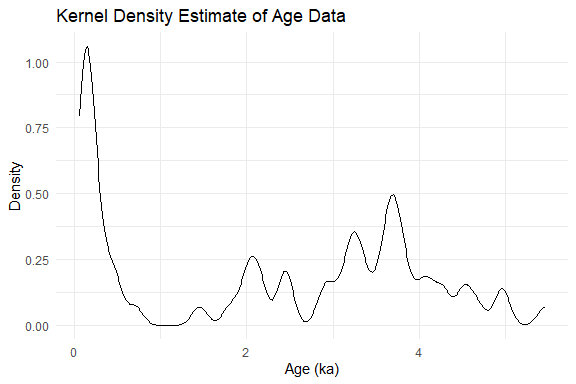
### Assessing the contribution of the last 0.5 ka

acrop <-raw\_ages|>  
 filter(source == "This paper")|>  
 filter(species\_2 == "BRAN")  
  
age\_acrop <- acrop$t\_ka  
# Calculate the bandwidth using unbiased cross-validation & constructing the KDE  
bw\_mx\_ucv <- bw.ucv(age\_acrop)

#> Warning in bw.ucv(age\_acrop): minimum occurred at one end of the range

kde\_plot <- ggplot(acrop, aes(x = t\_ka)) +  
 geom\_density(aes(y = ..density..), bw = bw\_mx\_ucv) +  
 labs(title = "Kernel Density Estimate of Age Data",  
 x = "Age (ka)",  
 y = "Density") +  
 theme\_minimal()  
  
print(kde\_plot)

#> Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.  
#> ℹ Please use `after\_stat(density)` instead.  
#> This warning is displayed once every 8 hours.  
#> Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
#> generated.



# Defining the KDE function  
kde\_function <- density(age\_acrop, bw = bw\_mx\_ucv)  
  
# Calculating the total area under the KDE curve  
total\_area <- sum(kde\_function$y) \* diff(kde\_function$x)[1]  
  
# Calculating the area for the range 0 to 0.5ka  
partial\_area <- sum(kde\_function$y[kde\_function$x >= 0 & kde\_function$x <= 0.7]) \* diff(kde\_function$x)[1]  
  
# Calculate last 0.5ka partial contribution  
partial\_contribution <- partial\_area / total\_area  
  
partial\_contribution

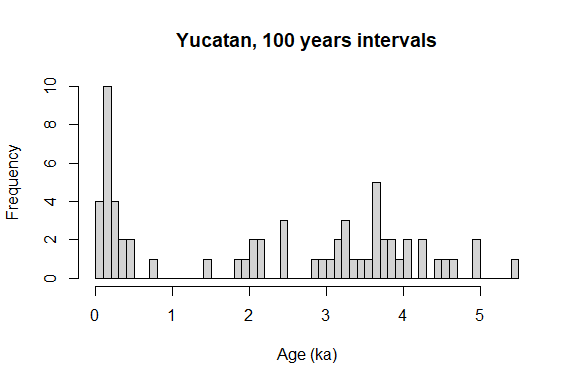
#> [1] 0.3282475

## Monte Carlo Simulation of Gap-size probability

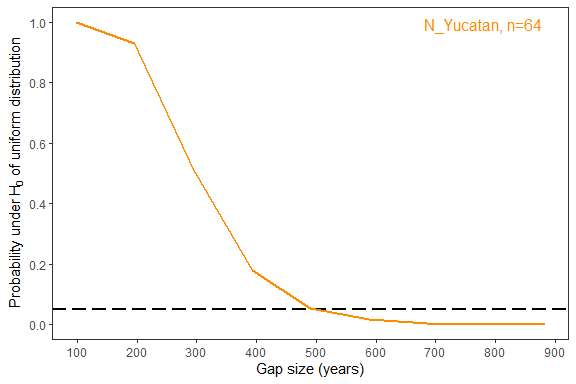
### Simulation using Uniform distribution

## Yucatan

age\_range <- 1000 \* (max(age\_acrop) - min(age\_acrop))  
  
intervals <- 100 #length of class intervals in years  
  
CI <- round(age\_range/intervals, 0) # number of classes  
  
# Histogram to visualize the data  
hist(age\_acrop,  
 breaks = CI,   
 main = paste("Yucatan,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")  
  
# Defining class intervals  
breaks <- hist(age\_acrop,   
 breaks = CI,   
 main = paste("Yucatan,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")$breaks



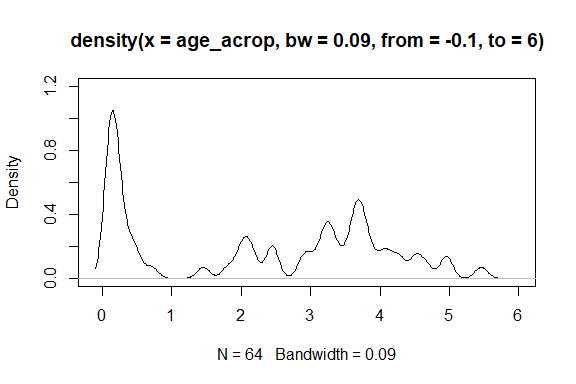
# Calculate the observed frequencies  
observed <- hist(age\_acrop, breaks = breaks, plot = FALSE)$counts  
  
# Function to identify sequences of zeros and their length  
gaps <- function(vector) {  
 secuencias <- data.frame(inicio = integer(), longitud = integer())  
 longitud\_actual <- 0  
 inicio\_secuencia <- NULL  
   
 for (i in 1:length(vector)) {  
 if (vector[i] == 0) {  
 if (is.null(inicio\_secuencia)) {  
 inicio\_secuencia <- i  
 }  
 longitud\_actual <- longitud\_actual + 1  
 } else {  
 if (!is.null(inicio\_secuencia)) {  
 secuencias <- rbind(secuencias, data.frame(inicio = inicio\_secuencia, longitud = longitud\_actual))  
 longitud\_actual <- 0  
 inicio\_secuencia <- NULL  
 }  
 }  
 }  
   
 if (!is.null(inicio\_secuencia)) {  
 secuencias <- rbind(secuencias, data.frame(inicio = inicio\_secuencia, longitud = longitud\_actual))  
 }  
   
 secuencias$size <- secuencias$longitud \* (age\_range/length(observed))  
   
   
 return(secuencias)  
}  
  
# Monte Carlo Simulation  
n\_sim <- 10000  
simulated\_gaps <- numeric(n\_sim)  
simulated\_gaps\_size <- vector(mode = "list", length = n\_sim)  
  
for (i in 1:n\_sim) {  
 # Parameterized Uniform Distribution  
 simulated\_data <- runif(length(age\_acrop), min = min(age\_acrop), max = max(age\_acrop))  
   
 # Calculate the simulated frequencies  
 sim\_observed <- hist(simulated\_data, breaks = CI, plot = FALSE)$counts  
   
 # Gap tests  
 simulated\_gaps[i] <- sum(sim\_observed == 0)  
 simulated\_gaps\_size[[i]] <- gaps(sim\_observed)  
}  
  
# Gap size analysis  
gap\_size <- list\_rbind(simulated\_gaps\_size, names\_to = "simulation")|>  
 group\_by(simulation, size)|>  
 reframe(f = n())  
  
yucatan <- gap\_size|>  
 group\_by(size)|>  
 summarise(average = round(mean(f),0), times = length(size), freq = (times+1)/(n\_sim+1))|>  
 mutate(site = "Yucatan")  
  
# plotting gaps probability in response to gap size  
  
ggplot(yucatan, aes(x = size, y = freq)) +  
 scale\_y\_continuous(breaks = seq(0, 1, 0.2)) +  
 scale\_x\_continuous(breaks = seq(0, 1400, 100)) +  
 geom\_hline(yintercept = 0.05,  
 colour = "black",  
 linewidth = 1,   
 linetype = "longdash") +  
 ylab(expression("Probability under " \* H[0] \* " of uniform distribution")) +  
 xlab("Gap size (years)") +  
 geom\_line(aes(color = "N\_Yucatan, n=64"), linewidth = 1) +   
 scale\_color\_manual(values = c("N\_Yucatan, n=64" = "darkorange"), name = "Legend") +  
 theme\_bw() +  
 theme(  
 panel.grid.major.x = element\_blank(),  
 panel.grid.minor.x = element\_blank(),  
 panel.grid.minor.y = element\_blank(),  
 panel.grid.major.y = element\_blank(),  
 legend.position = "none") +  
 guides(color = guide\_legend(override.aes = list(linetype = 1, size = 1))) +  
 annotation\_custom(  
 grob = grid::grobTree(  
 grid::textGrob("N\_Yucatan, n=64", x = 0.95, y = 0.95, hjust = 1, gp = grid::gpar(col = "darkorange"))  
 ),  
 xmin = -Inf, xmax = Inf, ymin = -Inf, ymax = Inf  
 )



## Gap uncertainties & sample size effect

### Gap uncertainties analysis

#calculating density estimates for the apal ages  
pdf0<-density(age\_acrop,bw=0.09,from=-0.1,to=6)  
#ploting KDE   
plot(pdf0,xlim=c(-0.1,6),ylim=c(0,1.2))



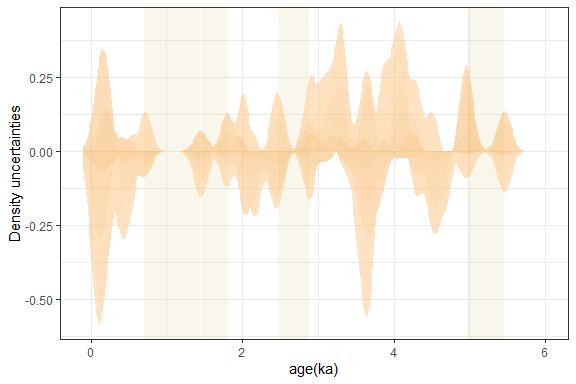
# randomly sampling test  
pdf25\_1<-density(age\_acrop[sample(1:51,25)],bw=0.09,from=-0.1,to=6)  
b<-pdf0$x[2]-pdf0$x[1]  
sum(abs(pdf0$y-pdf25\_1$y)\*b)/2

#> [1] 0.2285641

# Creating a function to generate densities with different sample sizes   
generar\_densidades <- function(datos, inicio = 36, fin = 64) {  
# Creating empty list to store results  
 lista\_densidades <- list()  
# Loop for each sample size  
 for(n in seq(inicio, fin, by = 4)) {  
# generating Sample   
 muestra <- datos[sample(1:length(datos), n)]  
# Calculating density  
 densidad <- density(muestra, bw = 0.09, from = -0.1, to = 6)  
# Storing densities   
 lista\_densidades[[paste0("pdf", n)]] <- densidad  
 }  
# return densities list  
 return(lista\_densidades)  
}  
# function use  
densidades <- generar\_densidades(age\_acrop)  
# empty dataframe  
df\_densidades <- data.frame()  
# density loop  
for(i in names(densidades)) {  
# temporal dataframe  
 df\_temp <- data.frame(x = densidades[[i]]$x, y = densidades[[i]]$y, group = i)  
# adding temporal to main dataframe  
 df\_densidades <- rbind(df\_densidades, df\_temp)  
}  
  
#Evaluating uncertainties in densities  
# Empty dataframe to store density differences  
df\_diferencias <- data.frame()  
  
# Creating a loop for each density on the list  
for(i in names(densidades)) {  
 # Creating temporal dataframe to store delta densities  
 df\_temp <- data.frame(x = pdf0$x, y = pdf0$y - densidades[[i]]$y, group = i)  
   
 # Adding temporal dataframe with delta densities to the main dataframe  
 df\_diferencias <- rbind(df\_diferencias, df\_temp)  
}  
  
#ploting gap uncertainties  
library(paletteer)

#> Warning: package 'paletteer' was built under R version 4.3.3

colores2<-paletteer\_c("ggthemes::Orange", 30)  
ggplot(df\_diferencias, aes(x = x, y = y, fill = group)) +  
 annotate("rect", xmin = 0.71, xmax = 1.44, ymin = -Inf, ymax = Inf, alpha = 0.15, fill = "#DAC778") +   
 annotate("rect", xmin = 1.46, xmax = 1.81, ymin = -Inf, ymax = Inf, alpha = 0.15, fill = "#DAC778") +  
 annotate("rect", xmin = 2.47, xmax = 2.88, ymin = -Inf, ymax = Inf, alpha = 0.15, fill = "#DAC778") +  
 annotate("rect", xmin = 4.97, xmax = 5.46, ymin = -Inf, ymax = Inf, alpha = 0.15, fill = "#DAC778")+  
 geom\_area(alpha = 0.5) +  
 scale\_fill\_manual(values = colores2) +  
 labs(x = "age(ka)", y = "Density uncertainties", fill = "Group") +  
 coord\_cartesian(xlim = c(-0.1, 6))+  
 theme\_bw()+  
 theme(legend.position = "none")

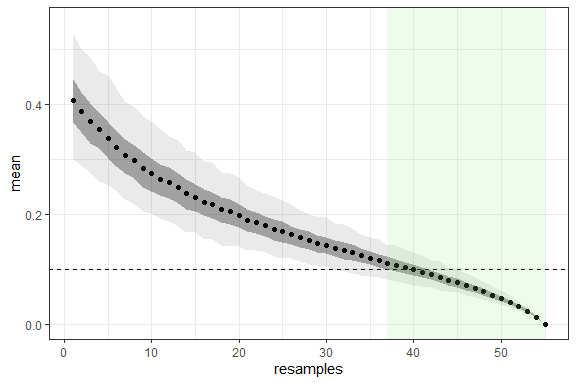


### Sample size analysis

#Accessing resampling related uncertainties  
desajuste<-matrix(ncol=1000,nrow=55)  
for(j in 1:nrow(desajuste)){  
 for(i in 1:ncol(desajuste)){  
 desajuste[j,i]<-sum(abs(pdf0$y -   
 density(age\_acrop[sample(1:64,(j+9))],bw=0.09,from=-0.1,to=6)$y)\*b)/2  
 }  
}  
  
dim(desajuste)

#> [1] 55 1000

# Create a new data frame for plotting PUMA age densities  
df <- data.frame(  
 x = 1:55,  
 mean = apply(desajuste, 1, mean),  
 q25 = apply(desajuste, 1, quantile, 0.25),  
 q75 = apply(desajuste, 1, quantile, 0.75),  
 q025 = apply(desajuste, 1, quantile, 0.025),  
 q975 = apply(desajuste, 1, quantile, 0.975)  
)  
  
# plotting sample size impact   
ggplot(df, aes(x = x, y = mean)) +  
 geom\_ribbon(aes(ymin = q025, ymax = q975), alpha = 0.1) +  
 geom\_ribbon(aes(ymin = q25, ymax = q75), alpha = 0.4) +  
 geom\_point(shape = 16) +  
 labs(x = "resamples", y = "mean") +  
 scale\_x\_continuous(breaks = seq(0, 60, by = 10)) +  
 scale\_y\_continuous(limits = c(0, 0.55)) +  
 theme\_bw()+  
 geom\_hline(yintercept = 0.1, color = "black", linetype = "dashed")+  
 annotate("rect", xmin = 37, xmax = 55, ymin = -Inf, ymax = Inf, alpha = 0.15, fill = "lightgreen")



## Regional trends (Exploratory analysis of published data for comparative purpouses)

regional<-read\_excel("raw\_ages.xlsx")# ADJUSTED TEMPORAL FRAME   
  
# regional figure  
fNW\_Car <- regional[regional$spot == "Northwestern Caribbean" , ] # This paper, Mexico  
fSW\_Car <- regional[regional$spot == "Southwestern Caribbean" , ] # Belize  
fE\_Car <- regional[regional$spot == "Eastern Caribbean" , ] # Saint Croix  
fLA\_Car <- regional[regional$spot == "Lesser Antilles"& regional$source =="Abdul et al.,16", ] # Barbados  
  
  
summary(as.factor(fNW\_Car$group))

#> Other Species palmata   
#> 27 64

summary(as.factor(fSW\_Car$group))

#> Other species palmata   
#> 52 69

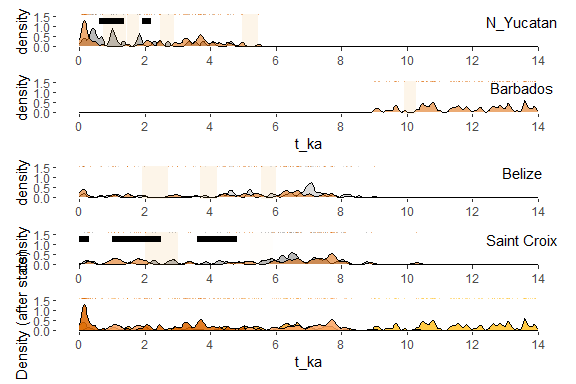
summary(as.factor(fE\_Car$group))

#> Other Species palmata   
#> 45 52

summary(as.factor(fLA\_Car$group))

#> palmata   
#> 81

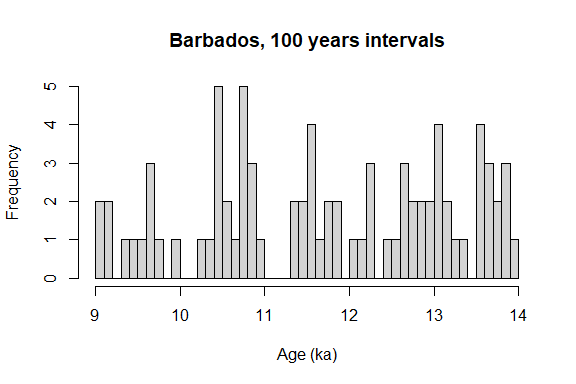
A<-ggplot(fNW\_Car) +  
 aes(x = t\_ka) +  
 geom\_density(aes(y = after\_stat(density), fill = group), alpha = 0.55, adjust = 0.09) +  
 scale\_fill\_manual(values = c("palmata" = "#dc6601","Other species" = "gray")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), legend.position = "none") +  
 scale\_x\_continuous(limits = c(0, 14), breaks = seq(0, 14, 2))+  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 1/2))+  
 theme(strip.background = element\_blank(), panel.background = element\_blank(),  
 legend.position = "none", axis.title.x = element\_blank())+  
 annotate("rect", xmin = 0.71, xmax = 1.44, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +   
 annotate("rect", xmin = 1.46, xmax = 1.81, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +  
 annotate("rect", xmin = 2.47, xmax = 2.88, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +  
 annotate("rect", xmin = 4.97, xmax = 5.46, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27")+  
 annotate("segment", x = 0.6, xend = 1.35, y = 1.25, yend = 1.25, color = "black", linewidth = 2.25) +  
 annotate("segment", x = 1.9, xend = 2.2, y = 1.25, yend = 1.25, color = "black", linewidth = 2.25) +  
 annotate("text", x = 13.5, y = 1.25, label = "N\_Yucatan", color="black")+  
 geom\_rug(aes(y = 1.5, colour = group), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray"))  
  
B<-ggplot(fSW\_Car) +  
 aes(x = t\_ka) +  
 geom\_density(aes(y = after\_stat(density), fill = group), alpha = 0.55, adjust = 0.088) +  
 scale\_fill\_manual(values = c("palmata" = "#dc6601","Other species" = "gray")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), legend.position = "none") +  
 scale\_x\_continuous(limits = c(0, 14), breaks = seq(0, 14, 2))+  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 1/2))+  
 theme(strip.background = element\_blank(), panel.background = element\_blank(),  
 legend.position = "none", axis.title.x = element\_blank())+  
 annotate("rect", xmin = 1.9, xmax = 2.72, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +   
 annotate("rect", xmin = 3.68, xmax = 4.2, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +  
 annotate("rect", xmin = 5.54, xmax = 6.0, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +  
 annotate("text", x = 13.5, y = 1.25, label = "Belize", color="black") +  
 geom\_rug(aes(y = 1.5, colour = group), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray"))  
  
C<-ggplot(fE\_Car) +  
 aes(x = t\_ka) +  
 geom\_density(aes(y = after\_stat(density), fill = group), alpha = 0.55, adjust = 0.088, linetype = "solid") +  
 scale\_fill\_manual(values = c("palmata" = "#dc6601","Other species" = "gray")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), axis.title.x = element\_blank(),legend.position = "none") +  
 scale\_x\_continuous(limits = c(0, 14), breaks = seq(0, 14, 2))+  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 1/2))+  
 annotate("rect", xmin = 2.0, xmax = 3.0, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +  
 annotate("rect", xmin = 5.2, xmax = 5.9, ymin = -Inf, ymax = Inf, alpha = 0.01, fill = "#ee9f27")+  
 annotate("segment", x = 3.6, xend = 4.8, y = 1.25, yend = 1.25, color = "black",linewidth= 2.25)+  
 annotate("segment", x = 1.0, xend = 2.5, y = 1.25, yend = 1.25, color = "black",linewidth= 2.25)+  
 annotate("segment", x = 0.0, xend = 0.3, y = 1.25, yend = 1.25, color = "black",linewidth= 2.25)+  
 annotate("text", x = 13.5, y = 1.25, label = "Saint Croix", color="black")+  
 geom\_rug(aes(y = 1.5, colour = group), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray"))  
  
D<-ggplot(fLA\_Car) +  
 aes(x = t\_ka) +  
 geom\_density(aes(y = after\_stat(density), fill = group), alpha = 0.55, adjust = 0.088) +  
 scale\_fill\_manual(values = c("palmata" = "#dc6601","Other species" = "gray")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), legend.position = "none") +  
 scale\_x\_continuous(limits = c(0,14), breaks = seq(0, 14, 2))+  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 1/2))+  
 annotate("rect", xmin = 9.92, xmax = 10.28, ymin = -Inf, ymax = Inf, alpha = 0.1, fill = "#ee9f27") +   
 annotate("text", x = 13.5, y = 1.25, label = "Barbados", color="black")+  
 geom\_rug(aes(y = 1.5, colour = group), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("palmata" = "#dc6601", "Other species" = "gray"))  
  
regional\_palmata <- subset(regional, group == "palmata")  
  
E<-ggplot(regional\_palmata, aes(x = t\_ka)) +  
 geom\_density(aes(y = after\_stat(density), fill = source), alpha = 0.85, adjust = 0.088) +  
 scale\_fill\_manual(values = c("This paper" = "#dc6601", "Gischler et al.,23" = "#ee9f27","Hubbard et al.,05;13" = "#FCA360", "Abdul et al.,16" = "#FFC125")) +  
 theme(strip.background = element\_blank(), panel.background = element\_blank(), , legend.position = "none") +  
 labs(x = "t\_ka", y = "Density (after stats)") +  
 scale\_x\_continuous(limits = c(0,14), breaks = seq(0, 14, 2))+  
 scale\_y\_continuous(limits = c(0, 1.5), breaks = seq(0, 1.5, 1/2))+  
 geom\_rug(aes(y = 1.5, colour = source), sides = "t", linewidth = 0.5, alpha = 0.5) +  
 scale\_colour\_manual(values = c("This paper" = "#dc6601", "Gischler et al.,23" = "#ee9f27","Hubbard et al.,05;13" = "#FCA360", "Abdul et al.,16" = "#FFC125"))  
  
combined\_plot <- A/D/B/C/E  
combined\_plot



## Monte Carlo Simulation of Gap-size probability (Regional Scope).

## Barbados

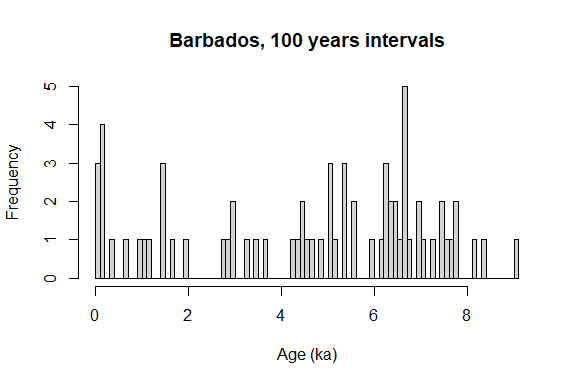
acrop <-raw\_ages|>  
 filter(source == "Abdul et al.,16")|>  
 filter(species\_2 == "BRAN")  
  
age\_acrop <- acrop$t\_ka  
  
age\_range <- 1000 \* (max(age\_acrop) - min(age\_acrop))  
  
intervals <- 100  
  
CI <- round(age\_range/intervals, 0)  
  
# Histogram to visualize the data  
hist(age\_acrop,  
 breaks = CI,   
 main = paste("Barbados,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")  
  
# Defining class intervals  
breaks <- hist(age\_acrop,   
 breaks = CI,   
 main = paste("Barbados,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")$breaks



# Calculate the observed frequencies  
observed <- hist(age\_acrop, breaks = breaks, plot = FALSE)$counts  
  
# Monte Carlo Simulation  
n\_sim <- 10000  
simulated\_gaps <- numeric(n\_sim)  
simulated\_gaps\_size <- vector(mode = "list", length = n\_sim)  
  
for (i in 1:n\_sim) {  
 simulated\_data <- runif(81, min = min(age\_acrop), max = max(age\_acrop))  
   
 sim\_observed <- hist(simulated\_data, breaks = CI, plot = FALSE)$counts  
   
 simulated\_gaps[i] <- sum(sim\_observed == 0)  
 simulated\_gaps\_size[[i]] <- gaps(sim\_observed)  
}  
  
# Gap size analysis  
  
gap\_size <- list\_rbind(simulated\_gaps\_size, names\_to = "simulation")|>  
 group\_by(simulation, size)|>  
 reframe(f = n())  
  
barbados <- gap\_size|>  
 group\_by(size)|>  
 summarise(average = round(mean(f),0), times = length(size), freq = (times+1)/(n\_sim+1))|>  
 mutate(site = "Barbados")

## Belize

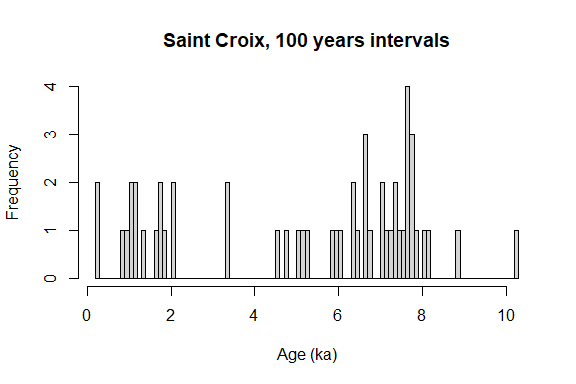
acrop <-raw\_ages|>  
 filter(source == "Gischler et al.,23")|>  
 filter(species\_2 == "BRAN")  
  
age\_acrop <- acrop$t\_ka  
  
age\_range <- 1000 \* (max(age\_acrop) - min(age\_acrop))  
  
intervals <-100  
  
CI <- round(age\_range/intervals, 0)  
  
# Histogram to visualize the data  
hist(age\_acrop,  
 breaks = CI,   
 main = paste("Barbados,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")  
  
# Defining class intervals  
breaks <- hist(age\_acrop,   
 breaks = CI,   
 main = paste("Barbados,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")$breaks



# Calculate the observed frequencies  
observed <- hist(age\_acrop, breaks = breaks, plot = FALSE)$counts  
  
# Monte Carlo simulation  
n\_sim <- 10000  
simulated\_gaps <- numeric(n\_sim)  
simulated\_gaps\_size <- vector(mode = "list", length = n\_sim)  
  
for (i in 1:n\_sim) {  
 simulated\_data <- runif(69, min = min(age\_acrop), max = max(age\_acrop))  
   
 sim\_observed <- hist(simulated\_data, breaks = CI, plot = FALSE)$counts  
   
 simulated\_gaps[i] <- sum(sim\_observed == 0)  
 simulated\_gaps\_size[[i]] <- gaps(sim\_observed)  
}  
  
# Gap size analysis  
  
gap\_size <- list\_rbind(simulated\_gaps\_size, names\_to = "simulation")|>  
 group\_by(simulation, size)|>  
 reframe(f = n())  
  
belize <- gap\_size|>  
 group\_by(size)|>  
 summarise(average = round(mean(f),0), times = length(size), freq = (times+1)/(n\_sim+1))|>  
 mutate(site = "Belize")

## St. Croix

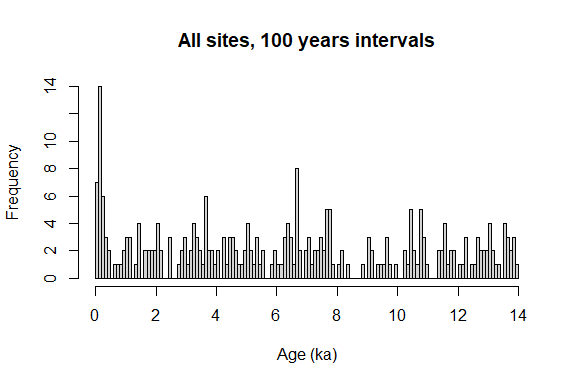
acrop <-raw\_ages|>  
 filter(source == "Hubbard et al.,05;13")|>  
 filter(species\_2 == "BRAN")  
  
age\_acrop <- acrop$t\_ka  
  
age\_range <- 1000 \* (max(age\_acrop) - min(age\_acrop))  
  
intervals <- 100  
  
CI <- round(age\_range/intervals, 0)  
  
# Histogram to visualize the data  
hist(age\_acrop,  
 breaks = CI,   
 main = paste("Saint Croix,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")  
  
# Defining class intervals  
breaks <- hist(age\_acrop,   
 breaks = CI,   
 main = paste("Saint Croix,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")$breaks



# Calculate the observed frequencies  
observed <- hist(age\_acrop, breaks = breaks, plot = FALSE)$counts  
  
# Monte Carlo simulations  
n\_sim <- 10000  
simulated\_gaps <- numeric(n\_sim)  
simulated\_gaps\_size <- vector(mode = "list", length = n\_sim)  
  
for (i in 1:n\_sim) {  
 simulated\_data <- runif(52, min = min(age\_acrop), max = max(age\_acrop))  
   
 sim\_observed <- hist(simulated\_data, breaks = CI, plot = FALSE)$counts  
   
 simulated\_gaps[i] <- sum(sim\_observed == 0)  
 simulated\_gaps\_size[[i]] <- gaps(sim\_observed)  
}  
  
# Gap size analysis  
gap\_size <- list\_rbind(simulated\_gaps\_size, names\_to = "simulation")|>  
 group\_by(simulation, size)|>  
 reframe(f = n())  
  
stc <- gap\_size|>  
 group\_by(size)|>  
 summarise(average = round(mean(f),0), times = length(size), freq = (times+1)/(n\_sim+1))|>  
 mutate(site = "Saint Croix")

### Regional scope for large gap probabilities

raw\_ages <- read\_excel("raw\_ages.xlsx")  
acrop <-raw\_ages|>  
 filter(species\_2 == "BRAN")  
  
age\_acrop <- acrop$t\_ka  
  
age\_range <- 1000 \* (max(age\_acrop) - min(age\_acrop))  
  
intervals <- 100  
  
CI <- round(age\_range/intervals, 0)  
  
# Histogram to visualize the data  
hist(age\_acrop,  
 breaks = CI,   
 main = paste("All sites,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")  
  
# Defining class intervals  
breaks <- hist(age\_acrop,   
 breaks = CI,   
 main = paste("All sites,", intervals, "years intervals"),  
 xlab = "Age (ka)",  
 ylab = "Frequency")$breaks



# Calculate the observed frequencies  
observed <- hist(age\_acrop, breaks = breaks, plot = FALSE)$counts  
  
# Monte Carlo simulation  
n\_sim <- 10000  
simulated\_gaps <- numeric(n\_sim)  
simulated\_gaps\_size <- vector(mode = "list", length = n\_sim)  
  
for (i in 1:n\_sim) {  
 simulated\_data <- runif(266, min = min(age\_acrop), max = max(age\_acrop))  
   
 sim\_observed <- hist(simulated\_data, breaks = CI, plot = FALSE)$counts  
   
 simulated\_gaps[i] <- sum(sim\_observed == 0)  
 simulated\_gaps\_size[[i]] <- gaps(sim\_observed)  
}  
  
# Gap size analysis  
  
gap\_size <- list\_rbind(simulated\_gaps\_size, names\_to = "simulation")|>  
 group\_by(simulation, size)|>  
 reframe(f = n())  
  
all\_sites <- gap\_size|>  
 group\_by(size)|>  
 summarise(average = round(mean(f),0), times = length(size), freq = (times+1)/(n\_sim+1))|>  
 mutate(site = "All sites")  
  
  
bind\_rows(yucatan, barbados, belize, stc, all\_sites) |>  
 filter(size <= 1300) |>  
 mutate(Region = factor(  
 site,  
 levels = c("Yucatan", "Barbados", "Belize", "Saint Croix", "All sites"),  
 labels = c(  
 "Yucatan (n = 64)",  
 "Barbados (n = 81)",  
 "Belize (n = 69)",  
 "Saint Croix (n = 52)",  
 "All sites (n = 266)"  
 )  
 )) |>  
 ggplot(aes(x = size, y = freq)) +  
 scale\_y\_continuous(breaks = seq(0, 1, 0.2)) +  
 scale\_x\_continuous(breaks = seq(50, 14000, 100)) +  
 geom\_hline(yintercept = 0.05,  
 colour = "black",  
 linetype="longdash",  
 linewidth = 1) +  
 ylab(expression("Probability under " \* H[0] \* " of uniform distribution") ) +  
 xlab("Gap size (years)") +  
 geom\_line(aes(colour = Region), linewidth = 1) +  
 theme\_bw() +  
 theme(  
 panel.grid.major.x = element\_blank(),  
 panel.grid.minor.x = element\_blank(),  
 panel.grid.minor.y = element\_blank(),  
 panel.grid.major.y = element\_blank(),  
 legend.position = "inside",  
 legend.justification = c(0.95, 0.95),  
 legend.background = element\_rect(fill = "white", colour = "white")  
 )

