

Script1

Fernanda Gonzalez-Saldias

2025-06-02

Contents

1	Set working space	1
1.1	Load libraries	1
1.2	Load morphological and molecular of relative abundances dataset	2
2	Procrustes analysis	3
2.1	Distances calculation	3
2.2	Procrustes between morphological to molecular identification	3
2.3	Create the residual figure	3
2.4	Procrustes calculation	4
2.5	Display Figure 3	6
3	Lineal model	7
3.1	Load metadata, alpha diversity calculation and Statistical analysis	7
3.2	To make a new data frame with pH mean, minimum and max for location and alpha index .	11
3.3	Models and figures	12
3.4	Display Figure 4	24
4	Correlation and log2 fold changes	25
4.1	Log2 fold changes	25
4.2	Correlation analysis	30
4.3	Display Figure 2	32

1 Set working space

1.1 Load libraries

```
rm(list=ls())
library(vegan)
library(ggplot2)
library(dplyr)
library(readxl)
library(xlsx)
library(Hmisc)
library(ggcorrplot)
library(ggthemes)
library(corrplot)
library(vegan)
library(cowplot)
library(car)
```

```
library("pgirmess")
library(microbiome)
```

1.2 Load morphological and molecular of relative abundances dataset

```
Morpho=data.frame(read_xlsx("data\\Table S2.xlsx",sheet="Morphological",na=""))
head(Morpho)[,1:5]
```

##	Station	Location	pH	Achnanthidium.minutisimum	Achnanthidium.caledonicum
## 1	BAI01	BAI	5.45	1.4	0
## 2	BAI04	BAI	4.95	0.0	0
## 3	BAI06	BAI	5.05	0.0	0
## 4	BAI07	BAI	5.10	8.8	0
## 5	BPU04	BPU	6.22	2.8	0
## 6	BSC03	BSC	7.55	6.4	0

```
COI=data.frame(read_xlsx("data\\Table S2.xlsx",sheet="COI",na=""))
head(COI)[,1:5]
```

##	Station	Location	pH	BOGC_000020776	BOGC_000020780
## 1	BAI01	BAI	5.45	0	0
## 2	BAI04	BAI	4.95	0	0
## 3	BAI06	BAI	5.05	0	0
## 4	BAI07	BAI	5.10	0	0
## 5	BPU04	BPU	6.22	0	0
## 6	BSC03	BSC	7.55	0	0

```
X18s=data.frame(read_xlsx("data\\Table S2.xlsx",sheet="18S rRNA",na=""))
head(X18s)[,1:5]
```

##	Station	Location	pH	BOGS_000001575	BOGS_000014580
## 1	BAI01	BAI	5.45	0	0.000000
## 2	BAI04	BAI	4.95	0	0.000000
## 3	BAI06	BAI	5.05	0	0.000000
## 4	BAI07	BAI	5.10	0	0.000000
## 5	BPU04	BPU	6.22	0	0.000000
## 6	BSC03	BSC	7.55	0	1.886792

#assigned rownames

```
rownames(Morpho)=(Morpho$Station)
rownames(COI)=(COI$Station)
rownames(X18s)=(X18s$Station)
```

#Select the numeric data

```
dato_morpho<-Morpho[4:386]
dato_coi<-COI[4:280]#277 otus
dato_18s<-X18s[4:169]#166 otus
```

#square root transformation of relative abundance data

```
datos.trans_morpho <- sqrt(dato_morpho)
datos.trans_coi <- sqrt(dato_coi)
datos.trans_18s<- sqrt(dato_18s)
```

2 Procrustes analysis

2.1 Distances calculation

```
dist.BC_morpho <- vegdist(datos.trans_morpho)
dist.BC_coi <- vegdist(datos.trans_coi)
dist.BC_18s <- vegdist(datos.trans_18s)

#distances
out.mds_morpho <- cmdscale(dist.BC_morpho,eig=T,k=2)
out.mds_coi <- cmdscale(dist.BC_coi,eig=T,k=2)
out.mds_18s <- cmdscale(dist.BC_18s ,eig=T,k=2)
```

2.2 Procrustes between morphological to molecular identification

```
#Residual calculation
#COI
pro_morpho_coi<- procrustes(X = out.mds_morpho,
                           Y = out.mds_coi, symmetric = TRUE,choices = c(1,2))
Res.coi=data.frame(resid=residuals(pro_morpho_coi),
                  Station=names(residuals(pro_morpho_coi)),Method="COI")

#18S
pro_morpho_18s<- procrustes(X = out.mds_morpho,
                           Y = out.mds_18s, symmetric = TRUE)
Res.18s=data.frame(resid=residuals(pro_morpho_18s),
                  Station=names(residuals(pro_morpho_18s)),Method="18S")

#Concatenated residual of procrustes dataframe
Res.all<-rbind(Res.coi,Res.18s)

#a merge is performed with the data matrix to add the location, station and pH.
Res.all=merge(Res.all,Morpho[,1:3],by="Station",sort = FALSE)

#Set method as a factor
Res.all$Method=as.factor(Res.all$Method)

#reorder factors
Res.all$Method=factor(Res.all$Method,c("COI","18S"))
head(Res.all)
```

```
##      Station      resid Method Location    pH
## 1   BAI01 0.02872557   COI      BAI 5.45
## 2   BAI01 0.05415680   18S      BAI 5.45
## 3   BAI04 0.03054378   18S      BAI 4.95
## 4   BAI04 0.05504238   COI      BAI 4.95
## 5   BAI06 0.02213755   18S      BAI 5.05
## 6   BAI06 0.09839896   COI      BAI 5.05
```

2.3 Create the residual figure

```
#Display panel figure 3C
Residuals=Res.all$>%
```

```

ggplot(aes(Method,resid))+
geom_boxplot(outlier.shape = NA)+
geom_jitter(width = 0.2,shape=21,size=2)+labs(x=" ",y="Residuals",title = "")+
ylim(0,0.18)+
theme_bw()+ theme(strip.text.x = element_text(size=12),
                  strip.background = element_blank(),
                  panel.grid.minor = element_blank(),
                  panel.grid.major = element_blank(),
                  axis.text.x = element_text(size=12,color="black",
                                              angle=0,vjust = 0.5),
                  axis.text.y = element_text(size=8,color="black",
                                              angle=0,vjust = 0.5),
                  axis.title.x = element_text(color="black", size=10,
                                              margin = margin(t = 5)),
                  axis.title.y = element_text(color="black", size=10),
                  text = element_text(family="serif"))

```

#Normality test

```
tapply(Res.all$resid,Res.all$Method,shapiro.test)
```

```

## $COI
##
##  Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.93771, p-value = 0.002481
##
##
## $`18S`
##
##  Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.94212, p-value = 0.004018

```

#Statistical test

```
wilcox.test(Res.all$resid~Res.all$Method)
```

```

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  Res.all$resid by Res.all$Method
## W = 2345, p-value = 0.4486
## alternative hypothesis: true location shift is not equal to 0

```

2.4 Procrustes calculation

#Morphological v/s COI procrustes calculation

```

pro_morpho_coi<- procrustes(X = out.mds_morpho,
                           Y = out.mds_coi, symmetric = TRUE,choices = c(1,2))

```

#statistical test

```

protest(X = out.mds_morpho,
        Y = out.mds_coi, scores = "sites", permutations = 999)

```

```
##
```

```

## Call:
## protest(X = out.mds_morpho, Y = out.mds_coi, scores = "sites",      permutations = 999)
##
## Procrustes Sum of Squares (m12 squared):      0.3974
## Correlation in a symmetric Procrustes rotation: 0.7763
## Significance: 0.001
##
## Permutation: free
## Number of permutations: 999

#Get data from pro_morfo_coi
tmp.pro<-rbind(pro_morpho_coi[["Yrot"]],pro_morpho_coi[["X"]])

pro.df<-data.frame(X=tmp.pro[,1],
                  Y=tmp.pro[,2],
                  Method=rep(c("Molecular","Morphological"),each=66),
                  Station=row.names(tmp.pro))

pro.df$Method<-as.factor(pro.df$Method)
pro.df$Method<-factor(pro.df$Method, c("Molecular","Morphological"))

pro.df<-merge(pro.df,COI[,1:3],by="Station",sort = FALSE)

#Display panel figure 3A
pro_morpho_coi=pro.df%>%
  ggplot(aes(X,Y,color=pH,shape=Method))+
  geom_line(aes(group=Station),color="black")+
  geom_point(size=4)+
  theme_bw()+labs(x="Dimension 1",y="Dimension 2",title = "Morphological vs. COI")+
  theme(strip.text.x = element_text(size=12),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        plot.title = element_text(size=12,hjust = 0.5),
        axis.title.x =element_text(size=10,color="black"),
        axis.title.y= element_text(size=10,color="black"),
        axis.text.x = element_text(size=8,color="black",angle=0,vjust = 0.5),
        axis.text.y = element_text(size=8,color="black",angle=0,vjust = 0.5),
        text = element_text(family="serif"))+
  scale_color_gradientn(colours = c("#F27F0C","yellow","#7ed348","#26B170","#01377d"))+
  theme(legend.position = "none")

#Morphological v/s 18S procrustes calculation
pro_morpho_18s<- procrustes(X = out.mds_morpho, Y = out.mds_18s, symmetric = TRUE)
#Statistical test
protest(X = out.mds_morpho, Y = out.mds_18s, scores = "sites", permutations = 999)

##
## Call:
## protest(X = out.mds_morpho, Y = out.mds_18s, scores = "sites",      permutations = 999)
##
## Procrustes Sum of Squares (m12 squared):      0.317
## Correlation in a symmetric Procrustes rotation: 0.8264
## Significance: 0.001
##
## Permutation: free
## Number of permutations: 999

```

```

#Get data from pro_morfo_18s
tmp.pro_2<-rbind(pro_morpha_18s[["Yrot"]],pro_morpha_18s[["X"]])

pro.df_2<-data.frame(X=tmp.pro_2[,1],
                    Y=tmp.pro_2[,2],
                    Method=rep(c("Molecular","Morphological"),each=66),
                    Station=row.names(tmp.pro_2))

pro.df_2$Method<-as.factor(pro.df_2$Method)
pro.df_2$Method<-factor(pro.df_2$Method, c("Molecular","Morphological"))

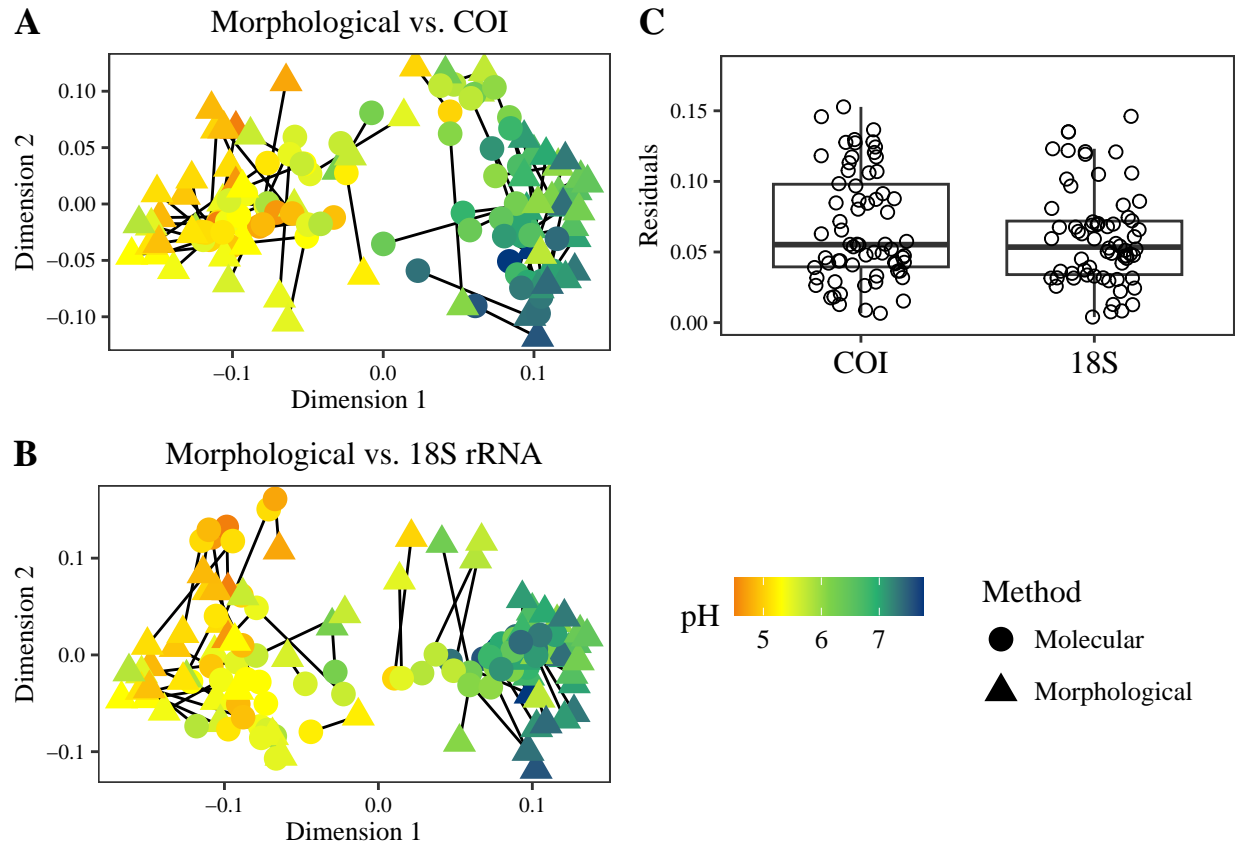
pro.df_2<-merge(pro.df_2,X18s[,1:3],by="Station",sort = FALSE)

#Display panel figure 3B
pro_morpha_18s=pro.df_2%>%
  ggplot(aes(X,Y,color=pH,shape=Method))+
  geom_line(aes(group=Station),color="black")+
  geom_point(size=4)+
  theme_bw()+labs(x="Dimension 1",y="Dimension 2",title = "Morphological vs. 18S rRNA")+
  theme(strip.text.x = element_text(size=12),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        plot.title = element_text(size=12,hjust = 0.5),
        axis.title.x =element_text(size=10,color="black"),
        axis.title.y= element_text(size=10,color="black"),
        axis.text.x = element_text(size=8,color="black",angle=0,vjust = 0.5),
        axis.text.y = element_text(size=8,color="black",angle=0,vjust = 0.5),
        text = element_text(family="serif"))+
  scale_color_gradientn(colours = c("#F27F0C","yellow","#7ed348","#26B170","#01377d"))+
  theme(legend.position = "bottom",legend.key.size = unit(0.5, 'cm'),
        legend.text = element_text(size=10),
        legend.title = element_text(size = 12))+
  guides(shape = guide_legend(title.position = "top",ncol=1))
legenda_18s_pro=get_plot_component(pro_morpha_18s, 'guide-box-bottom', return_all = TRUE)
pro_morpha_18s=pro_morpha_18s+theme(legend.position = "none")

```

2.5 Display Figure 3

Figure 3: Comparison of MDS through Procrustes analysis between morphological vs. COI (A) and morphological vs. 18S (B). Residual values from comparing morphological identification and the two molecular identifications (18S and COI) (C).



3 Lineal model

3.1 Load metadata, alpha diversity calculation and Statistical analysis

```
#Load metadata
metadata=data.frame(read_xlsx("data\\Table S1.xlsx",sheet="Metadata",na=""))

#alpha diversity calculation
metadata$H_morpho=vegan::diversity(t(t(Morpho[,4:386])))
metadata$S_morpho=vegan::specnumber(t(t(Morpho[,4:386])))
metadata$J_morpho=evenness((t(Morpho[,4:386])), 'pielou')[,1]

metadata$H_COI=vegan::diversity(t(t(COI[,4:280])))
metadata$S_COI=vegan::specnumber(t(t(COI[,4:280])))
metadata$J_COI=evenness((t(COI[,4:280])), 'pielou')[,1]

metadata$H_18s=vegan::diversity(t(t(X18s[,4:169])))
metadata$S_18s=vegan::specnumber(t(t(X18s[,4:169])))
metadata$J_18s=evenness((t(X18s[,4:169])), 'pielou')[,1]
head(metadata)[,1:5]
```

```
## Station Location WTD GDD Rad
## 1 BAI01 BAI -1.0 1532 1835
## 2 BAI04 BAI -0.5 1410 1745
## 3 BAI06 BAI -13.0 1532 1835
```

```

## 4   BAI07      BAI   -1.5 1593 1742
## 5   BPU04      BPU   -1.0 2903 1812
## 6   BSC03      BSC   -2.0 2906 1946

morpho_index=metadata[,1:2]
morpho_index$Type<-"Morpho"
morpho_index$Richness=metadata[,25]
morpho_index$Shannon=metadata[,24]
morpho_index$Pielou=metadata[,26]

coi_index=metadata[,1:2]
coi_index$Type<-"COI"
coi_index$Richness=metadata[,28]
coi_index$Shannon=metadata[,27]
coi_index$Pielou=metadata[,29]

x18s_index=metadata[,1:2]
x18s_index$Type<-"18S"
x18s_index$Richness=metadata[,31]
x18s_index$Shannon=metadata[,30]
x18s_index$Pielou=metadata[,32]

analisis=rbind(morpho_index,coi_index,x18s_index)
head(analisis)[,1:5]

##   Station Location   Type Richness  Shannon
## 1   BAI01      BAI Morpho      68 3.458610
## 2   BAI04      BAI Morpho      40 2.797134
## 3   BAI06      BAI Morpho      43 3.062109
## 4   BAI07      BAI Morpho      61 3.528707
## 5   BPU04      BPU Morpho      61 3.346099
## 6   BSC03      BSC Morpho      35 2.823989

#Richness
#Normality test
tapply(analisis$Richness,analisis$Type,shapiro.test)

## $`18S`
##
##   Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.97336, p-value = 0.1664
##
##
## $COI
##
##   Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.9613, p-value = 0.03773
##
##
## $Morpho
##

```



```

## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.9878, p-value = 0.7653
#homoscedasticity test
bartlett.test(analysis$Richness~analysis$Type)

##
## Bartlett test of homogeneity of variances
##
## data:  analysis$Richness by analysis$Type
## Bartlett's K-squared = 46.003, df = 2, p-value = 1.024e-10
#kruskal Wallis test calculation
kruskal_S=kruskal.test(analysis$Richness~analysis$Type)
kruskal_S

##
## Kruskal-Wallis rank sum test
##
## data:  analysis$Richness by analysis$Type
## Kruskal-Wallis chi-squared = 131.8, df = 2, p-value < 2.2e-16
#Post test to evaluate where are the significant differences
kruskalmc(analysis$Richness~analysis$Type,alpha=0.001)

## Multiple comparison test after Kruskal-Wallis
## alpha: 0.001
## Comparisons
##      obs.dif critical.dif stat.signif
## 18S-COI      37.28030      35.78934      TRUE
## 18S-Morpho 112.35606      35.78934      TRUE
## COI-Morpho  75.07576      35.78934      TRUE
#Shannon index
#Normality test
tapply(analysis$Shannon,analysis$Type,shapiro.test)

## $`18S`
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.94625, p-value = 0.006386
##
##
## $COI
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.95526, p-value = 0.0182
##
##
## $Morpho
##

```

```

## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.94707, p-value = 0.007016
#homoscedasticity test
bartlett.test(analysis$Shannon~analysis$Type)

##
## Bartlett test of homogeneity of variances
##
## data:  analysis$Shannon by analysis$Type
## Bartlett's K-squared = 18.125, df = 2, p-value = 0.000116
#kruskal Wallis test calculation
kruskal_H=kruskal.test(analysis$Shannon~analysis$Type)
kruskal_H

##
## Kruskal-Wallis rank sum test
##
## data:  analysis$Shannon by analysis$Type
## Kruskal-Wallis chi-squared = 75.169, df = 2, p-value < 2.2e-16
#Post test to evaluate where are the significant differences
kruskalmc(analysis$Shannon~analysis$Type,alpha=0.01)

## Multiple comparison test after Kruskal-Wallis
## alpha: 0.01
## Comparisons
##      obs.dif critical.dif stat.signif
## 18S-COI    30.75758      29.27852      TRUE
## 18S-Morpho  85.37879      29.27852      TRUE
## COI-Morpho  54.62121      29.27852      TRUE
#Pielou evenness
#Normality test
tapply(analysis$Pielou,analysis$Type,shapiro.test)

## $`18S`
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.86626, p-value = 3.992e-06
##
##
## $COI
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.83768, p-value = 5.165e-07
##
##
## $Morpho
##

```

```
## Shapiro-Wilk normality test
##
## data: X[[i]]
## W = 0.89811, p-value = 5.31e-05

#homoscedasticity test
bartlett.test(analisis$Pielou~analisis$Type)

##
## Bartlett test of homogeneity of variances
##
## data: analisis$Pielou by analisis$Type
## Bartlett's K-squared = 34.65, df = 2, p-value = 2.991e-08

#kruskal Wallis test calculation
kruskal_J=kruskal.test(analisis$Pielou~analisis$Type)
kruskal_J

##
## Kruskal-Wallis rank sum test
##
## data: analisis$Pielou by analisis$Type
## Kruskal-Wallis chi-squared = 1.2121, df = 2, p-value = 0.5455
```

3.2 To make a new data frame with pH mean, minimum and max for location and alpha index

```
ind_ph=aggregate(pH~Location,data=metadata,mean)
ind_ph$pH.sd=aggregate(pH~Location,data=metadata,sd)[,2]
ind_ph$S_morpho=aggregate(S_morpho~Location,data=metadata,mean)[,2]
ind_ph$S_morpho.sd=aggregate(S_morpho~Location,data=metadata,sd)[,2]
ind_ph$S_COI=aggregate(S_COI~Location,data=metadata,mean)[,2]
ind_ph$S_COI.sd=aggregate(S_COI~Location,data=metadata,sd)[,2]
ind_ph$S_18S=aggregate(S_18s~Location,data=metadata,mean)[,2]
ind_ph$S_18S.sd=aggregate(S_18s~Location,data=metadata,sd)[,2]

ind_ph$H_morpho=aggregate(H_morpho~Location,data=metadata,mean)[,2]
ind_ph$H_morpho.sd=aggregate(H_morpho~Location,data=metadata,sd)[,2]
ind_ph$H_COI=aggregate(H_COI~Location,data=metadata,mean)[,2]
ind_ph$H_COI.sd=aggregate(H_COI~Location,data=metadata,sd)[,2]
ind_ph$H_18S=aggregate(H_18s~Location,data=metadata,mean)[,2]
ind_ph$H_18S.sd=aggregate(H_18s~Location,data=metadata,sd)[,2]

ind_ph$J_morpho=aggregate(J_morpho~Location,data=metadata,mean)[,2]
ind_ph$J_morpho.sd=aggregate(J_morpho~Location,data=metadata,sd)[,2]
ind_ph$J_COI=aggregate(J_COI~Location,data=metadata,mean)[,2]
ind_ph$J_COI.sd=aggregate(J_COI~Location,data=metadata,sd)[,2]
ind_ph$J_18S=aggregate(J_18s~Location,data=metadata,mean)[,2]
ind_ph$J_18S.sd=aggregate(J_18s~Location,data=metadata,sd)[,2]
head(ind_ph)

## Location      pH      pH.sd S_morpho S_morpho.sd      S_COI S_COI.sd
## 1 BAI 5.137500 0.21746647 53.00000 13.638182 16.75000 5.852350
## 2 BPU 6.220000 NA 61.00000 NA 18.00000 NA
## 3 BSC 7.556667 0.20008332 36.66667 10.598742 15.66667 11.930353
```

```
## 4      BV 5.360000 0.05656854 33.00000      8.485281 16.00000 4.242641
## 5     BValp 5.416667 0.50649120 41.00000      3.464102 12.33333 5.131601
## 6      CBR 5.425000 0.54447222 59.50000      4.949747 27.50000 3.535534
##      S_18S S_18S.sd H_morpho H_morpho.sd      H_COI H_COI.sd      H_18S H_18S.sd
## 1 11.000000 4.690416 3.211640      0.3443353 2.165513 0.3376578 1.434446 0.8080834
## 2  6.000000      NA 3.346099      NA 2.435441      NA 1.438600      NA
## 3 14.666667 2.081666 2.781214      0.5113690 1.411986 1.3421487 2.035147 0.1318731
## 4 12.000000 4.242641 2.776004      0.3641266 2.411507 0.3189316 1.712559 0.3619781
## 5  9.666667 4.041452 2.961465      0.1834044 1.643052 0.6098899 1.662429 0.4842821
## 6 12.500000 3.535534 3.256676      0.1061851 2.499575 0.2564536 2.064514 0.4135034
##      J_morpho J_morpho.sd      J_COI J_COI.sd      J_18S J_18S.sd
## 1 0.8126118 0.041235912 0.7806790 0.03414478 0.5893503 0.23983976
## 2 0.8139630      NA 0.8426046      NA 0.8028979      NA
## 3 0.7738577 0.082931258 0.4807212 0.35589246 0.7597906 0.03626807
## 4 0.7960840 0.045155725 0.8739245 0.03064695 0.6949203 0.04524402
## 5 0.7975268 0.032274672 0.6539655 0.14926927 0.7424420 0.08198796
## 6 0.7972762 0.009740795 0.7542136 0.04810213 0.8199782 0.07121246
```

3.3 Models and figures

3.3.1 Richness

```
#Morphological
lineal.poly=lm(S_morpho~poly(pH,2),data=ind_ph)
summary(lineal.poly)

##
## Call:
## lm(formula = S_morpho ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.475  -4.691  -2.249   5.421  15.993
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    42.163      1.667   25.298 <2e-16 ***
## poly(pH, 2)1    -2.615      8.498   -0.308  0.7611
## poly(pH, 2)2   -21.889      8.498   -2.576  0.0169 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.498 on 23 degrees of freedom
## Multiple R-squared:  0.2263, Adjusted R-squared:  0.1591
## F-statistic: 3.365 on 2 and 23 DF,  p-value: 0.05227

lineal.lm=lm(S_morpho~pH,data=ind_ph)
summary(lineal.lm)

##
## Call:
## lm(formula = S_morpho ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -21.493 -6.510 -1.288 7.270 18.946
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  46.1071    14.3602   3.211  0.00374 **
## pH          -0.6517     2.3531  -0.277  0.78420
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.443 on 24 degrees of freedom
## Multiple R-squared:  0.003185, Adjusted R-squared:  -0.03835
## F-statistic: 0.07669 on 1 and 24 DF, p-value: 0.7842

AIC(lineal.poly,lineal.lm)

##              df      AIC
## lineal.poly   4 189.8707
## lineal.lm     3 194.4601

#Display panel figure 4A
S_morpho=ind_ph%>%
  ggplot(aes(pH,S_morpho))+
  geom_point() +
  geom_errorbar(aes(ymin = S_morpho-S_morpho.sd,ymax = S_morpho+S_morpho.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x =element_blank(),axis.title.x= element_blank(),
        axis.text.y = element_text(size=10, color="black"),
        axis.title = element_text(color="black",size=12, face="bold"),
        text = element_text(family="serif"))+
  labs(y="Richness",x="pH", title = "      Morphological") +xlim(4,8)+ylim(0,70)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.8,y=11,
        label=bquote("y"~.(round(summary(lineal.poly)[[4]][3,1],3))~
          "x"~2~"+".(round(summary(lineal.poly)[[4]][2,1],3))~
          "x"~"+".(round(summary(lineal.poly)[[4]][1,1],3))),
        cex = 2.5,col="black")+
  annotate("text",x=5.8,y=5,
        label=bquote("R"~2~"="~.(round(summary(lineal.poly)[[8]],3))),
        cex = 2.5,col="black")

#COI
lineal.poly=lm(S_COI~poly(pH,2),data=ind_ph)
summary(lineal.poly)

##
## Call:
## lm(formula = S_COI ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.6806  -4.2949   0.0523   2.8785  12.6813
##
## Coefficients:
```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    20.469      1.200  17.063 1.49e-14 ***
## poly(pH, 2)1     9.995      6.117   1.634  0.1159
## poly(pH, 2)2   -21.579      6.117  -3.528  0.0018 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.117 on 23 degrees of freedom
## Multiple R-squared:  0.3966, Adjusted R-squared:  0.3441
## F-statistic: 7.557 on 2 and 23 DF,  p-value: 0.003002

lineal.lm=lm(S_COI~pH,data=ind_ph)
summary(lineal.lm)
```

```
##
## Call:
## lm(formula = S_COI ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.203  -5.627  -1.670   4.369  14.941
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.397      11.304   0.477   0.637
## pH              2.490       1.852   1.344   0.191
##
## Residual standard error: 7.434 on 24 degrees of freedom
## Multiple R-squared:  0.07004, Adjusted R-squared:  0.0313
## F-statistic: 1.808 on 1 and 24 DF,  p-value: 0.1914

AIC(lineal.poly,lineal.lm)
```

```
##           df      AIC
## lineal.poly  4 172.7735
## lineal.lm    3 182.0179
```

#Display panel figure 4B

```
S_COI=ind_ph$>%
  ggplot(aes(pH,S_COI))+
  geom_point() +
  geom_errorbar(aes(ymin = S_COI-S_COI.sd,ymax = S_COI+S_COI.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x =element_blank(),axis.title.x= element_blank(),
        axis.text.y =element_blank(),axis.title.y= element_blank(),
        text = element_text(family="serif"))+
  labs(y="",x="pH", title = "      COI") +xlim(4,8)+ylim(0,70)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.6,y=60,
        label=bquote("y=~.(round(summary(lineal.poly)[[4]][3,1],3))~
                      "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
                      "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex = 2.5,col="black")+
  
```

```

    annotate("text",x=5.6,y=55,
            label=bquote("R"~2~"="~.(round(summary(lineal.poly)[[8]],3))~"*"),
            cex = 2.5,col="black")

#18S rRNA
lineal.poly=lm(S_18S~poly(pH,2),data=ind_ph)
summary(lineal.poly)

##
## Call:
## lm(formula = S_18S ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.6486 -1.6226 -0.0661  0.6997  7.5010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   12.9353     0.6258  20.671 2.35e-16 ***
## poly(pH, 2)1    7.4697     3.1909   2.341  0.0283 *
## poly(pH, 2)2   -6.7718     3.1909  -2.122  0.0448 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.191 on 23 degrees of freedom
## Multiple R-squared:  0.3027, Adjusted R-squared:  0.2421
## F-statistic: 4.992 on 2 and 23 DF,  p-value: 0.01583

lineal.lm=lm(S_18S~pH,data=ind_ph)
summary(lineal.lm)

##
## Call:
## lm(formula = S_18S ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.2485 -1.9045 -0.7976  1.0671  8.6077
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.6713     5.1944   0.322  0.7504
## pH             1.8613     0.8512   2.187  0.0387 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.416 on 24 degrees of freedom
## Multiple R-squared:  0.1661, Adjusted R-squared:  0.1314
## F-statistic: 4.782 on 1 and 24 DF,  p-value: 0.03874

AIC(lineal.poly,lineal.lm)

##           df      AIC
## lineal.poly  4 138.9323
## lineal.lm    3 141.5820

```

```

#Display panel figure 4C
S_18s=ind_ph%>%
  ggplot(aes(pH,S_18S))+
  geom_point() +
  geom_errorbar(aes(ymin = S_18S-S_18S.sd,ymax = S_18S+S_18S.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x =element_blank(),axis.title.x= element_blank(),
        axis.text.y =element_blank(),axis.title.y= element_blank(),
        text = element_text(family="serif"))+
  labs(y="",x="pH", title = "      18S rRNA") +xlim(4,8)+ylim(0,70)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.5,y=60,
        label=bquote("y="~.(round(summary(lineal.poly)[[4]][3,1],3))~
          "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
          "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex = 2.5,col="black")+
  annotate("text",x=5.5,y=55,
        label=bquote("R"^2~"="~.(round(summary(lineal.poly)[[8]],3))~"*"),
        cex =2.5,col="black")

```

3.3.2 Shannon index

```

#Morphological
lineal.poly=lm(H_morpho~poly(pH,2),data=ind_ph)
summary(lineal.poly)

##
## Call:
## lm(formula = H_morpho ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32312 -0.18990  0.01622  0.19568  0.52735
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.76003    0.07813   35.328  <2e-16 ***
## poly(pH, 2)1  -0.38742    0.39836   -0.973    0.341
## poly(pH, 2)2  -0.76594    0.39836   -1.923    0.067 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3984 on 23 degrees of freedom
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0956
## F-statistic: 2.321 on 2 and 23 DF, p-value: 0.1207

lineal.lm=lm(H_morpho~pH,data=ind_ph)
summary(lineal.lm)

##
## Call:

```



```

## lm(formula = H_morpho ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39375 -0.13488  0.07533  0.25067  0.60231
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.34424    0.63890   5.234  2.3e-05 ***
## pH          -0.09654    0.10469  -0.922   0.366
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4201 on 24 degrees of freedom
## Multiple R-squared:  0.03422,    Adjusted R-squared:  -0.006025
## F-statistic: 0.8503 on 1 and 24 DF,  p-value: 0.3657
AIC(lineal.poly,lineal.lm)

##              df      AIC
## lineal.poly   4 30.73655
## lineal.lm     3 32.61192
#Display panel figure 4D
H_morpho=ind_ph%>%
  ggplot(aes(pH,H_morpho))+
  geom_point() +
  geom_errorbar(aes(ymin = H_morpho-H_morpho.sd,ymax = H_morpho+H_morpho.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x =element_blank(),axis.title.x= element_blank(),
        axis.text.y = element_text(size=10, color="black"),
        axis.title = element_text(color="black",size=12, face="bold"),
        text = element_text(family="serif"))+
  labs(y="Shannon",x="pH", title = "") +xlim(4,8)+ylim(0,4)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.6,y=1.0,
        label=bquote("y=~.(round(summary(lineal.poly)[[4]][3,1],3))~
          "x"^2~"+~.(round(summary(lineal.poly)[[4]][2,1],3))~
          "x"~"+~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex =2.5,col="black")+
  annotate("text",x=5.6,y=0.7,
        label=bquote("R"^2~"=~.(round(summary(lineal.poly)[[8]],3))),
        cex =2.5,col="black")

#COI
lineal.poly=lm(H_COI~poly(pH,2),data=ind_ph)
summary(lineal.poly)

```

```

##
## Call:
## lm(formula = H_COI ~ poly(pH, 2), data = ind_ph)
##
## Residuals:

```

```
##      Min      1Q   Median      3Q      Max
## -0.80267 -0.24028  0.04111  0.25047  0.54609
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.24431    0.07482  29.996 < 2e-16 ***
## poly(pH, 2)1  0.41430    0.38151   1.086   0.289
## poly(pH, 2)2 -2.03663    0.38151  -5.338 2.02e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3815 on 23 degrees of freedom
## Multiple R-squared:  0.5634, Adjusted R-squared:  0.5254
## F-statistic: 14.84 on 2 and 23 DF,  p-value: 7.264e-05
lineal.lm=lm(H_COI~pH,data=ind_ph)
summary(lineal.lm)
```

```
##
## Call:
## lm(formula = H_COI ~ pH, data = ind_ph)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.1302 -0.4353  0.1408  0.4006  0.8761
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.6196    0.8498   1.906  0.0687 .
## pH            0.1032    0.1393   0.741  0.4657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5588 on 24 degrees of freedom
## Multiple R-squared:  0.02239, Adjusted R-squared: -0.01835
## F-statistic: 0.5496 on 1 and 24 DF,  p-value: 0.4657
```

```
AIC(lineal.poly,lineal.lm)
```

```
##              df      AIC
## lineal.poly   4 28.48856
## lineal.lm     3 47.44596
```

```
#Display panel figure 4E
```

```
H_COI=ind_ph%>%
  ggplot(aes(pH,H_COI))+
  geom_point() +
  geom_errorbar(aes(ymin = H_COI-H_COI.sd,ymax = H_COI+H_COI.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x=element_blank(),axis.title.x= element_blank(),
        axis.text.y=element_blank(),axis.title.y= element_blank(),
        text = element_text(family="serif"))+
  labs(y="",x="pH", title = "") +xlim(4,8)+ylim(0,4)+
```

```
geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
annotate("text",x=5.5,y=3.9,
        label=bquote("y="~.(round(summary(lineal.poly)[[4]][3,1],3))~
                        "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
                        "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex =2.5,col="black")+
annotate("text",x=5.5,y=3.6,
        label=bquote("R"^2~"="~.(round(summary(lineal.poly)[[8]],3))~"***"),
        cex =2.5,col="black")
```

#18S rRNA

```
lineal.poly=lm(H_18S~poly(pH,2),data=ind_ph)
summary(lineal.poly)
```

```
##
## Call:
## lm(formula = H_18S ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.74800 -0.18582  0.00243  0.22456  0.81629
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.9150     0.0714  26.821 < 2e-16 ***
## poly(pH, 2)1    1.0767     0.3641   2.957  0.00706 **
## poly(pH, 2)2   -1.0954     0.3641  -3.009  0.00626 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3641 on 23 degrees of freedom
## Multiple R-squared:  0.4363, Adjusted R-squared:  0.3872
## F-statistic: 8.899 on 2 and 23 DF,  p-value: 0.001372
lineal.lm=lm(H_18S~pH,data=ind_ph)
summary(lineal.lm)
```

```
##
## Call:
## lm(formula = H_18S ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.77370 -0.25286 -0.07173  0.26940  0.99530
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2914     0.6398   0.455  0.6529
## pH             0.2683     0.1048   2.559  0.0172 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4207 on 24 degrees of freedom
## Multiple R-squared:  0.2144, Adjusted R-squared:  0.1816
## F-statistic: 6.549 on 1 and 24 DF,  p-value: 0.01722
```

```
AIC(lineal.poly,lineal.lm)
```

```
##           df      AIC
## lineal.poly  4 26.05427
## lineal.lm    3 32.68302
```

```
#Display panel figure 4F
```

```
H_18s=ind_ph%>%
  ggplot(aes(pH,H_18S))+
  geom_point() +
  geom_errorbar(aes(ymin = H_18S-H_18S.sd,ymax = H_18S+H_18S.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        #axis.text.x = element_text(size=10, color="black",hjust=1),
        #axis.title.x = element_text(color="black",size=12,face="bold"),
        axis.text.x =element_blank(),axis.title.x= element_blank(),
        axis.text.y =element_blank(),axis.title.y= element_blank(),
        text = element_text(family="serif"))+
  labs(y="",x="pH", title = "") +xlim(4,8)+ylim(0,4)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.5,y=3.8,
         label=bquote("y="~.(round(summary(lineal.poly)[[4]][3,1],3))~
                        "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
                        "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
         cex =2.5,col="black")+
  annotate("text",x=5.5,y=3.5,
         label=bquote("R"^2~"="~.(round(summary(lineal.poly)[[8]],3))~"*"),
         cex =2.5,col="black")
```

3.3.3 Pielou evenness

```
#Morphological
```

```
lineal.poly=lm(J_morpho~poly(pH,2),data=ind_ph)
summary(lineal.poly)
```

```
##
## Call:
## lm(formula = J_morpho ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.28127 -0.03504  0.01220  0.05196  0.12054
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.74030    0.01658  44.649  <2e-16 ***
## poly(pH, 2)1 -0.10425    0.08454  -1.233    0.230
## poly(pH, 2)2 -0.10164    0.08454  -1.202    0.241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08454 on 23 degrees of freedom
```

```
## Multiple R-squared:  0.1142, Adjusted R-squared:  0.03721
## F-statistic: 1.483 on 2 and 23 DF,  p-value: 0.2479

lineal.lm=lm(J_morpho~pH,data=ind_ph)
summary(lineal.lm)

##
## Call:
## lm(formula = J_morpho ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29065 -0.02830  0.02678  0.04397  0.11547
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.89751    0.12975   6.917 3.74e-07 ***
## pH          -0.02598    0.02126  -1.222   0.234
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08532 on 24 degrees of freedom
## Multiple R-squared:  0.05856,    Adjusted R-squared:  0.01934
## F-statistic: 1.493 on 1 and 24 DF,  p-value: 0.2336

AIC(lineal.poly,lineal.lm)

##              df      AIC
## lineal.poly   4 -49.8686
## lineal.lm     3 -50.2839

#Display panel figure 4G
J_morpho=ind_ph%>%
  ggplot(aes(pH,J_morpho))+
  geom_point() +
  geom_errorbar(aes(ymin = J_morpho-J_morpho.sd,ymax = J_morpho+J_morpho.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x = element_text(size=10, color="black",hjust=1),
        axis.title.x = element_text(color="black",size=12,face="bold"),
        axis.title = element_text(color="black",size=12, face="bold"),
        text = element_text(family="serif"))+
  labs(y="Pielou evenness",x="pH", title = "") +xlim(4,8)+ylim(0,1)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=5.6,y=0.32,
        label=bquote("y=~.(round(summary(lineal.poly)[[4]][3,1],3))~
                      "x"^2~"+~.(round(summary(lineal.poly)[[4]][2,1],3))~
                      "x"~"+~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex =2.5,col="black")+
  annotate("text",x=5.6,y=0.25,
        label=bquote("R"^2~"=~.(round(summary(lineal.poly)[[8]],3))),
        cex =2.5,col="black")
```

```

#COI
lineal.poly=lm(J_COI~poly(pH,2),data=ind_ph)
summary(lineal.poly)

##
## Call:
## lm(formula = J_COI ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.127062 -0.057129  0.004795  0.046659  0.169382
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.75152    0.01561  48.155 < 2e-16 ***
## poly(pH, 2)1 -0.01611    0.07958  -0.202   0.841
## poly(pH, 2)2 -0.43867    0.07958  -5.513 1.32e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07958 on 23 degrees of freedom
## Multiple R-squared:  0.5695, Adjusted R-squared:  0.5321
## F-statistic: 15.21 on 2 and 23 DF,  p-value: 6.171e-05

lineal.lm=lm(J_COI~pH,data=ind_ph)
summary(lineal.lm)

##
## Call:
## lm(formula = J_COI ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26476 -0.07588  0.02782  0.09231  0.13270
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.775819    0.180484   4.299 0.000247 ***
## pH            -0.004015    0.029575  -0.136 0.893148
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1187 on 24 degrees of freedom
## Multiple R-squared:  0.0007673, Adjusted R-squared:  -0.04087
## F-statistic: 0.01843 on 1 and 24 DF,  p-value: 0.8931

AIC(lineal.poly,lineal.lm)

##              df      AIC
## lineal.poly   4 -53.01649
## lineal.lm     3 -33.12181

#Display panel figure 4H
J_COI=ind_ph%>%
  ggplot(aes(pH,J_COI))+
  geom_point() +

```

```

geom_errorbar(aes(ymin = J_COI-J_COI.sd,ymax = J_COI+J_COI.sd )) +
geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
theme_bw()+
theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
      panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
      axis.text.x = element_text(size=10, color="black",hjust=1),
      axis.title.x = element_text(color="black",size=12,face="bold"),
      axis.text.y =element_blank(),axis.title.y= element_blank(),
      axis.title = element_text(color="black",size=12, face="bold"),
      text = element_text(family="serif"))+
labs(y="",x="pH", title = "") +xlim(4,8)+ylim(0,1)+
geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
annotate("text",x=5.8,y=0.32,
        label=bquote("y="~.(round(summary(lineal.poly)[[4]][3,1],3))~
          "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
          "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex = 2.5,col="black")+
annotate("text",x=5.8,y=0.25,
        label=bquote("R"^2~"="~.(round(summary(lineal.poly)[[8]],3))~"***"),
        cex =2.5,col="black")

#18S rRNA
lineal.poly=lm(J_18S~poly(pH,2),data=ind_ph)
summary(lineal.poly)

```

```

##
## Call:
## lm(formula = J_18S ~ poly(pH, 2), data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.190008 -0.034590 -0.002687  0.057936  0.143541
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.75325     0.01674  45.007 < 2e-16 ***
## poly(pH, 2)1   0.28137     0.08534   3.297 0.003152 **
## poly(pH, 2)2  -0.32960     0.08534  -3.862 0.000791 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08534 on 23 degrees of freedom
## Multiple R-squared:  0.5286, Adjusted R-squared:  0.4876
## F-statistic: 12.89 on 2 and 23 DF,  p-value: 0.0001755

lineal.lm=lm(J_18S~pH,data=ind_ph)
summary(lineal.lm)

```

```

##
## Call:
## lm(formula = J_18S ~ pH, data = ind_ph)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.270101 -0.086528  0.009426  0.061715  0.197407

```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.32896    0.16311   2.017  0.0550 .
## pH           0.07011    0.02673   2.623  0.0149 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1073 on 24 degrees of freedom
## Multiple R-squared:  0.2228, Adjusted R-squared:  0.1904
## F-statistic: 6.881 on 1 and 24 DF,  p-value: 0.0149

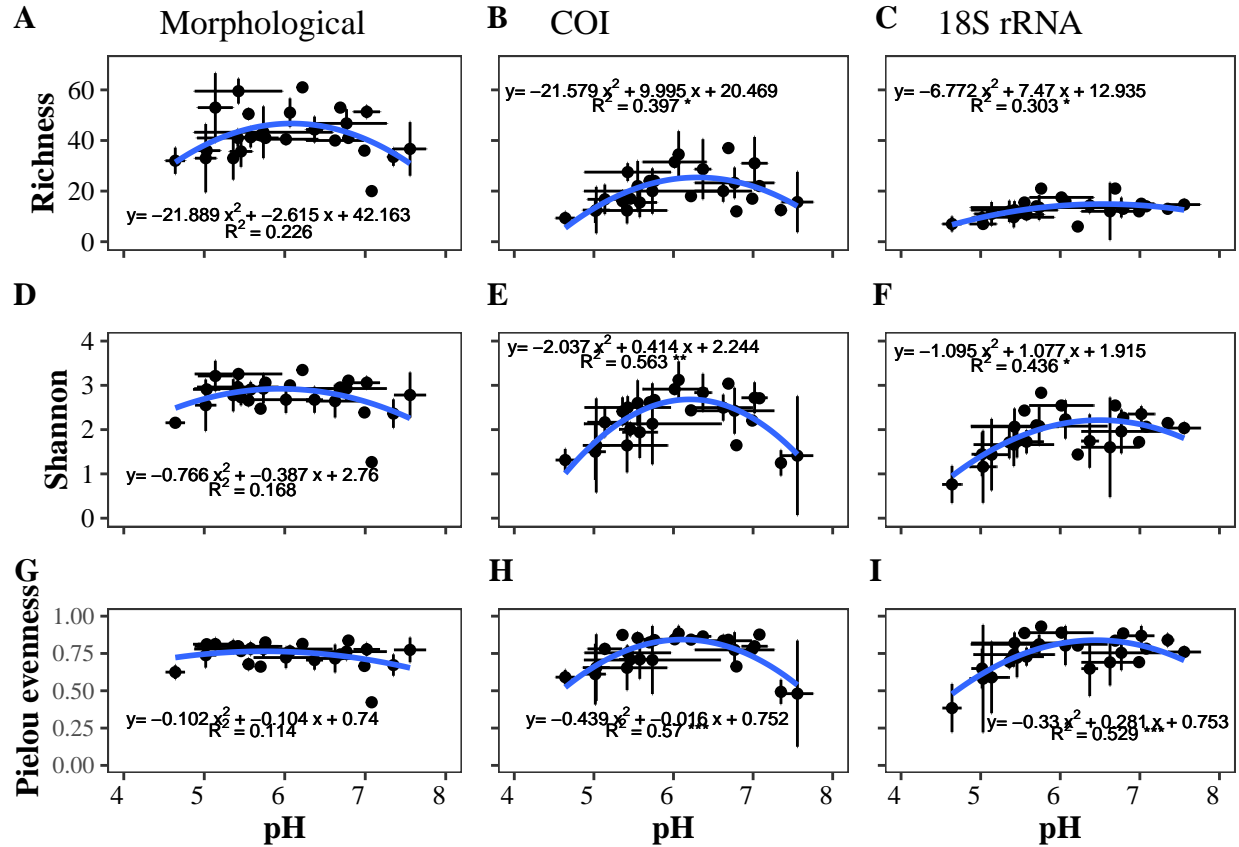
AIC(lineal.poly,lineal.lm)

##           df           AIC
## lineal.poly  4 -49.38161
## lineal.lm    3 -38.38386

#Display panel figure 4I
J_18s=ind_ph%>%
  ggplot(aes(pH,J_18S))+
  geom_point() +
  geom_errorbar(aes(ymin = J_18S-J_18S.sd,ymax = J_18S+J_18S.sd )) +
  geom_errorbarh(aes(xmin = pH-pH.sd,xmax = pH+pH.sd))+
  theme_bw()+
  theme(strip.text.x = element_text(size=10),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.text.x = element_text(size=10, color="black",hjust=1),
        axis.title.x = element_text(color="black",size=12,face="bold"),
        axis.text.y =element_blank(),axis.title.y= element_blank(),
        axis.title = element_text(color="black",size=12, face="bold"),
        text = element_text(family="serif"))+
  labs(y="",x="pH", title = "") +xlim(4,8)+ylim(0,1)+
  geom_smooth(se=FALSE,method="lm", formula=y~poly(x,2))+
  annotate("text",x=6.6,y=0.30,
        label=bquote("y=~.(round(summary(lineal.poly)[[4]][3,1],3))~
          "x"^2~"+"~.(round(summary(lineal.poly)[[4]][2,1],3))~
          "x"~"+"~.(round(summary(lineal.poly)[[4]][1,1],3))),
        cex = 2.5,col="black")+
  annotate("text",x=6.6,y=0.23,
        label=bquote("R"^2~"="~.(round(summary(lineal.poly)[[8]],3))~"***"),
        cex =2.5,col="black")
```

3.4 Display Figure 4

Figure 4: Correlation between pH and alpha diversity variables (Richness, Shannon index, and Pielou) through morphological (A, D, G), COI (B, E, H), and 18S rRNA (C, F, I) datasets ordered by increasing pH. The dot is the average pH vs the average of the corresponding alpha diversity variable of each locality. Lines are the standard deviations of pH on the horizontal axis and the corresponding alpha on the vertical axis. The blue line is an adjusted polynomial model of order two.



4 Correlation and log2 fold changes

4.1 Log2 fold changes

```
meta.diff.coi<-metadata[,1:2]
meta.diff.18s<-metadata[,1:2]

#Richnnes log2 fold changes
meta.diff.coi$L2FC<-log2(metadata$S_COI/metadata$S_morpho)
meta.diff.coi$Method<-"COI"
meta.diff.18s$L2FC<-log2(metadata$S_18s/metadata$S_morpho)
meta.diff.18s$Method<-"18S"

meta.diff.rich<-rbind(meta.diff.coi,meta.diff.18s)
head(meta.diff.rich)
```

##	Station	Location	L2FC	Method
## 1	BAI01	BAI	-1.502500	COI
## 2	BAI04	BAI	-1.415037	COI
## 3	BAI06	BAI	-2.104337	COI
## 4	BAI07	BAI	-1.760812	COI
## 5	BPU04	BPU	-1.760812	COI
## 6	BSC03	BSC	-1.544321	COI

```
#Display panel figure 2B
plot.L2FC.rich<-meta.diff.rich%>%
```

```

ggplot(aes(L2FC,Method,fill=Method))+
geom_boxplot(outlier.shape = NA,alpha=0.4)+
geom_jitter(aes(L2FC,Method,fill=Method),position=position_jitterdodge(0.4),
            shape=21,size=1)+
scale_fill_manual(values = c("#6600CC","#FF9900"))+geom_vline(xintercept = 0,
                                                            linetype="dashed")+

labs(x="",y="Richness")+xlim(-3.5,2)+
theme_bw()+theme(strip.text.x = element_text(size=10),
                  strip.background = element_blank(),
                  panel.grid.minor = element_blank(),
                  panel.grid.major = element_blank(),
                  axis.text.x =element_blank(),
                  axis.title.x= element_blank(),
                  axis.text.y = element_text(size=8,color="black",
                                                angle=0,vjust = 0.5),
                  axis.title.y = element_text(color="black", size=10),
                  text = element_text(family="serif"))+
  theme(legend.position = "right",legend.key.size = unit(0.5, 'cm'),
        legend.text = element_text(size=8))
legenda_rich=get_legend(plot.L2FC.rich)
plot.L2FC.rich=plot.L2FC.rich+theme(legend.position = "none")

#Normality test
shapiro.test(meta.diff.rich$L2FC[meta.diff.rich$Method=="COI"])

##
##  Shapiro-Wilk normality test
##
## data:  meta.diff.rich$L2FC[meta.diff.rich$Method == "COI"]
## W = 0.99045, p-value = 0.8952

shapiro.test(meta.diff.rich$L2FC[meta.diff.rich$Method=="18S"])

##
##  Shapiro-Wilk normality test
##
## data:  meta.diff.rich$L2FC[meta.diff.rich$Method == "18S"]
## W = 0.96967, p-value = 0.1057

#t-student
t.test(meta.diff.rich$L2FC[meta.diff.rich$Method=="COI"],mu=0)

##
##  One Sample t-test
##
## data:  meta.diff.rich$L2FC[meta.diff.rich$Method == "COI"]
## t = -16.661, df = 65, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -1.352556 -1.063000
## sample estimates:
## mean of x
## -1.207778

t.test(meta.diff.rich$L2FC[meta.diff.rich$Method=="18S"],mu=0)

```

```
##
## One Sample t-test
##
## data: meta.diff.rich$L2FC[meta.diff.rich$Method == "18S"]
## t = -25.065, df = 65, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -1.958868 -1.669743
## sample estimates:
## mean of x
## -1.814306
```

```
#Shannon index log2 fold changes
```

```
meta.diff.sha.coi<-metadata[,1:2]
```

```
meta.diff.sha.18s<-metadata[,1:2]
```

```
meta.diff.sha.coi$L2FC<-log2(metadata$H_COI/metadata$H_morpho)
```

```
meta.diff.sha.coi$Method<-"COI"
```

```
meta.diff.sha.18s$L2FC<-log2(metadata$H_18s/metadata$H_morpho)
```

```
meta.diff.sha.18s$Method<-"18S"
```

```
meta.diff.shannon<-rbind(meta.diff.sha.coi,meta.diff.sha.18s)
```

```
head(meta.diff.shannon)
```

```
## Station Location      L2FC Method
## 1  BAI01          BAI -0.4055937  COI
## 2  BAI04          BAI -0.3950673  COI
## 3  BAI06          BAI -0.7746882  COI
## 4  BAI07          BAI -0.7254976  COI
## 5  BPU04          BPU -0.4582974  COI
## 6  BSC03          BSC -1.3970570  COI
```

```
#Display panel figure 2D
```

```
plot.L2FC.shannon<-meta.diff.shannon%>%
```

```
  ggplot(aes(L2FC,Method,fill=Method))+
```

```
  geom_boxplot(outlier.shape = NA,alpha=0.4)+
```

```
  geom_jitter(aes(L2FC,Method,fill=Method),position=position_jitterdodge(0.4),
              shape=21,size=1)+
```

```
  scale_fill_manual(values = c("#6600CC","#FF9900"))+geom_vline(xintercept = 0,
                                                                linetype="dashed")+
```

```
  labs(x="",y="Shannon")+xlim(-3.5,2)+
```

```
  theme_bw()+theme(strip.text.x = element_text(size=10),
```

```
                    strip.background = element_blank(),
```

```
                    panel.grid.minor = element_blank(),
```

```
                    panel.grid.major = element_blank(),
```

```
                    axis.text.x =element_blank(),
```

```
                    axis.title.x= element_blank(),
```

```
                    axis.text.y = element_text(size=8,color="black",
                                                  angle=0,vjust = 0.5),
```

```
                    axis.title.y = element_text(color="black", size=10),
```

```
                    text = element_text(family="serif"))+
```

```
  theme(legend.position = "right",legend.key.size = unit(0.5, 'cm'),
```

```
        legend.text = element_text(size=8))
```

```
legenda_shannon=get_legend(plot.L2FC.shannon)
```

```
plot.L2FC.shannon=plot.L2FC.shannon+theme(legend.position = "none")
```

```

#Normality test
shapiro.test(meta.diff.shannon$L2FC[meta.diff.shannon$Method=="COI"])

##
## Shapiro-Wilk normality test
##
## data: meta.diff.shannon$L2FC[meta.diff.shannon$Method == "COI"]
## W = 0.91502, p-value = 0.0002481

shapiro.test(meta.diff.shannon$L2FC[meta.diff.shannon$Method=="18S"])

##
## Shapiro-Wilk normality test
##
## data: meta.diff.shannon$L2FC[meta.diff.shannon$Method == "18S"]
## W = 0.89594, p-value = 4.397e-05

#Statistical test
wilcox.test(meta.diff.shannon$L2FC[meta.diff.shannon$Method=="COI"],mu=0)

##
## Wilcoxon signed rank test with continuity correction
##
## data: meta.diff.shannon$L2FC[meta.diff.shannon$Method == "COI"]
## V = 242, p-value = 3.529e-08
## alternative hypothesis: true location is not equal to 0

wilcox.test(meta.diff.shannon$L2FC[meta.diff.shannon$Method=="18S"],mu=0)

##
## Wilcoxon signed rank test with continuity correction
##
## data: meta.diff.shannon$L2FC[meta.diff.shannon$Method == "18S"]
## V = 44, p-value = 1.221e-11
## alternative hypothesis: true location is not equal to 0

#Pielou evenness log2 fold changes
meta.diff.pie.coi<-metadata[,1:2]
meta.diff.pie.18s<-metadata[,1:2]

meta.diff.pie.coi$L2FC<-log2(metadata$J_COI/metadata$J_morpho)
meta.diff.pie.coi$Method<-"COI"
meta.diff.pie.18s$L2FC<-log2(metadata$J_18s/metadata$J_morpho)
meta.diff.pie.18s$Method<-"18S"

meta.diff.pielou<-rbind(meta.diff.pie.coi,meta.diff.pie.18s)
head(meta.diff.pielou)

## Station Location L2FC Method
## 1 BAI01 BAI 0.003337448 COI
## 2 BAI04 BAI 0.050860836 COI
## 3 BAI06 BAI -0.066749608 COI
## 4 BAI07 BAI -0.217307570 COI
## 5 BPU04 BPU 0.049892634 COI
## 6 BSC03 BSC -0.880257895 COI

```

```

#Display panel figure 2F
plot.L2FC.pielou<-meta.diff.pielou%>%
  ggplot(aes(L2FC,Method, fill=Method))+
  geom_boxplot(outlier.shape = NA,alpha=0.4)+
  geom_jitter(aes(L2FC,Method,fill=Method),position=position_jitterdodge(0.4),
              shape=21,size=1)+
  scale_fill_manual(values = c("#6600CC","#FF9900"))+geom_vline(xintercept = 0,
                                                                linetype="dashed")+
  labs(x=expression("Log" [2] * " fold change"),y="Pielou evenness")+xlim(-3.5,2)+
  theme_bw()+theme(strip.text.x = element_text(size=10),
                  strip.background = element_blank(),
                  panel.grid.minor = element_blank(),
                  panel.grid.major = element_blank(),
                  axis.text.x = element_text(size=8,color="black",
                                              angle=0,vjust = 0.5),
                  axis.text.y = element_text(size=8,color="black",
                                              angle=0,vjust = 0.5),
                  axis.title.x = element_text(color="black", size=10,
                                              margin = margin(t = 0,l=0)),
                  axis.title.y = element_text(color="black", size=10),
                  text = element_text(family="serif"))+
  guides(fill=guide_legend(ncol=2))+
  theme(legend.position="none",legend.key.size = unit(0.4, 'cm'),
        legend.text = element_text(size=4),legend.title = element_text(size=6))
legenda_pielou=get_legend(plot.L2FC.pielou)

```

#Normality test

```
shapiro.test(meta.diff.pielou$L2FC[meta.diff.pielou$Method=="COI"])
```

```

##
##  Shapiro-Wilk normality test
##
## data:  meta.diff.pielou$L2FC[meta.diff.pielou$Method == "COI"]
## W = 0.84996, p-value = 1.21e-06

```

```
shapiro.test(meta.diff.pielou$L2FC[meta.diff.pielou$Method=="18S"])
```

```

##
##  Shapiro-Wilk normality test
##
## data:  meta.diff.pielou$L2FC[meta.diff.pielou$Method == "18S"]
## W = 0.85743, p-value = 2.072e-06

```

#Statistical test

```
wilcox.test(meta.diff.pielou$L2FC[meta.diff.pielou$Method=="COI"],mu=0)
```

```

##
##  Wilcoxon signed rank test with continuity correction
##
## data:  meta.diff.pielou$L2FC[meta.diff.pielou$Method == "COI"]
## V = 1052, p-value = 0.7349
## alternative hypothesis: true location is not equal to 0
wilcox.test(meta.diff.pielou$L2FC[meta.diff.pielou$Method=="18S"],mu=0)

```

```
##
```

```
## Wilcoxon signed rank test with continuity correction
##
## data: meta.diff.pielou$L2FC[meta.diff.pielou$Method == "18S"]
## V = 1029, p-value = 0.6273
## alternative hypothesis: true location is not equal to 0
```

4.2 Correlation analysis

#Richness

```
corr_rich<-rcorr(as.matrix(metadata[,c(25,28,31)]),type="spearman")
corr_rich
```

```
##           S_morpho S_COI S_18s
## S_morpho      1.00  0.58  0.32
## S_COI         0.58  1.00  0.56
## S_18s         0.32  0.56  1.00
##
## n= 66
##
##
## P
##           S_morpho S_COI S_18s
## S_morpho           0.0000 0.0083
## S_COI       0.0000           0.0000
## S_18s       0.0083  0.0000
```

#significance

```
round(as.dist(corr_rich$P),4)
```

```
##           S_morpho S_COI
## S_COI      0.0000
## S_18s      0.0083 0.0000
```

#Display panel figure 2A

```
richness_corr=ggcorrplot(corr_rich$r,type = "upper",
                          outline.col = "white",lab=TRUE,insig = "blank",lab_size = 3)+
  theme_bw()+
  labs(x="",y="")+
  theme_tufte(base_size = 7)+
  theme(strip.text.x = element_text(size=12),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.title.x =element_blank(),axis.title.y= element_blank(),
        axis.text.x = element_text(size=8,color="black",angle=0,vjust = 0.5),
        axis.text.y = element_text(size=8,color="black",angle=0,vjust = 0.5),
        text = element_text(family="serif"))+
  theme(legend.position = "none",legend.key.size = unit(0.5, 'cm'),
        legend.text = element_text(size=8))
```

#Shannon index

```
corr_sha<-rcorr(as.matrix(metadata[,c(24,27,30)]),type="spearman")
corr_sha
```

```
##           H_morpho H_COI H_18s
## H_morpho      1.00  0.20  0.26
## H_COI         0.20  1.00  0.49
## H_18s         0.26  0.49  1.00
```

```
##
## n= 66
##
##
## P
##           H_morpho H_COI  H_18s
## H_morpho           0.1119 0.0375
## H_COI      0.1119           0.0000
## H_18s      0.0375    0.0000

#significance
round(as.dist(corr_sha$P),4)

##           H_morpho  H_COI
## H_COI      0.1119
## H_18s      0.0375 0.0000

#Display panel figure 2C
shannon_corr=ggcorrplot(corr_sha$r,type = "upper",
                        outline.col = "white",lab=TRUE,insig = "blank",lab_size = 3)+
  labs(x="",y="")+
  theme_bw()+
  theme_tufte(base_size = 7)+
  theme(strip.text.x = element_text(size=12),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.title.x =element_blank(),axis.title.y= element_blank(),
        axis.text.x = element_text(size=8,color="black",angle=0,vjust = 0.5),
        axis.text.y = element_text(size=8,color="black",angle=0,vjust = 0.5),
        text = element_text(family="serif"))+
  theme(legend.position = "none",legend.key.size = unit(0.5, 'cm'),
        legend.text = element_text(size=8))

#Pielou evenness
corr_pie<-rcorr(as.matrix(metadata[,c(26,29,32)]),type="spearman")
corr_pie

##           J_morpho J_COI J_18s
## J_morpho      1.00 -0.04  0.17
## J_COI        -0.04  1.00  0.23
## J_18s         0.17  0.23  1.00
##
## n= 66
##
##
## P
##           J_morpho J_COI  J_18s
## J_morpho           0.7376 0.1771
## J_COI      0.7376           0.0673
## J_18s      0.1771    0.0673

#Significance
round(as.dist(corr_pie$P),4)

##           J_morpho  J_COI
## J_COI      0.7376
## J_18s      0.1771 0.0673
```

```

#Display panel figure 2E
pielou_corr=ggcorrplot(corr_pie$r,type = "upper",
                      outline.col = "white",lab=TRUE,insig = "blank",lab_size=3)+
  labs(x="",y="")+
  theme_bw()+
  theme_tufte(base_size = 7)+
  theme(strip.text.x = element_text(size=12),strip.background = element_blank(),
        panel.grid.minor = element_blank(),panel.grid.major = element_blank(),
        axis.title.x =element_blank(),axis.title.y= element_blank(),
        axis.text.x = element_text(size=8,color="black",angle=0,vjust = 0.5),
        axis.text.y = element_text(size=8,color="black",angle=0,vjust = 0.5),
        text = element_text(family="serif"))+
  theme(legend.position = "bottom",legend.key.size = unit(0.4, 'cm'),
        legend.text = element_text(size=8),
        legend.title = element_text(size = 8))
legenda_pielou_corr=get_plot_component(pielou_corr, 'guide-box-bottom',
                                     return_all = TRUE)
pielou_corr=pielou_corr+theme(legend.position = "none")

```

4.3 Display Figure 2

Figure 2: Correlation between molecular and morphological identification in richness (A), Shannon index (C), and Pielou evenness (E). Comparison of alpha diversity indices in the three approaches to diatom communities of high mountain mires in richness (B), Shannon index (D), and Pielou evenness (F), molecular with morphological identification.

