

# seagrass diversity data analyses :

Ornstein Uhlenbeck process to simulate gene number shared by two species

J Chen

Last update: 2025-11-17

## Contents

Chapter 1. Life history traits vs genetic diversity	2
Chapter 2. DFE	6
Chapter 3. Effects of distribution range, Nc and historical Ne	22

Define some functions

```
# some functions
# sort dataset with rownames matching tree tip labels, required for phylo analyses
sort_data_by_treeTips=function(species, tree){
  tip.lab=tree$tip.label
  n=length(tip.lab)
  index=numeric(n)
  for(i in 1:n){
    index[i] = which(species == tip.lab[i])
  }
  return(index)
}

ci.lines=function(x, model, log=F, ...){
  x=x[!is.na(x)]
  xm<-mean(x)
  n<-length(x)
  ssx<-sum(x^2)-sum(x)^2/n
  s.t<-qt(0.975, (n-2))
  xv.min=min(x)
  xv.max=max(x)
  if(xv.min<0){
    xv.min=xv.min*1.2
  }else{
    xv.min=xv.min*0.6
  }
  if(xv.max<0){
    xv.max=xv.max*0.6
  }else{
    xv.max=xv.max*1.2
  }
  xv<-seq(xv.min, xv.max, length=100)
  yv<-coef(model)[1]+coef(model)[2]*xv
```

```

se<-sqrt(summary(model)[[6]]^2*(1/n+(xv-xm)^2/ssx))
ci<-s.t*se
uyv<-yv+ci
lyv<-yv-ci
if(log=='xy'){
  lines(10^xv,10^uyv, ...)
  lines(10^xv,10^lyv, ...)
}else if(log=='x'){
  lines(10^xv,uyv, ...)
  lines(10^xv,lyv, ...)
}
else if(log=='y'){
  lines(xv,10^uyv, ...)
  lines(xv,10^lyv, ...)
}
else{
  lines(xv,uyv, ...)
  lines(xv,lyv, ...)
}
}

plot_ci=function(xv, yv.lw, yv.hi, col = 1 , ...){
  n = length(xv)
  for(i in 1:n){
    arrows(xv[i],yv.lw[i], xv[i], yv.hi[i], angle=90, length = 0.06, code=3, col=col[i], ...)
  }
}

```

## Chapter 1. Life history traits vs genetic diversity

We first examined the distribution of genetic diversity across LHT

```

# process the data
data_family_daf=as.data.frame(read_excel('~/Desktop/ZJU/seagrass/1-manuscripts/202511/supplementaryTable1.xlsx'))
# classify into only two groups: 1) aerial & 2) hydrophilous
data_family_daf$pollination2 = ifelse(data_family_daf$pollination=='Hydrophilous', 'Hydrophilous', 'Aerial')
data_family_daf$distSites[data_family_daf$distSites<5]=NA
data_family_daf$observations[is.na(data_family_daf$distSites)]=NA
data_family_daf$census = data_family_daf$distSites/data_family_daf$leafSize
row.names(data_family_daf)= data_family_daf$Species

## remove Spirodela Polyrhiza as it is highly homozygous, very low diversity and possibly all cloned.
data_family_daf = data_family_daf[data_family_daf$Species!='SpirodelaPolyrhiza', ] #
tree<-read.tree('~/Desktop/ZJU/seagrass/1-manuscripts/202511/supplementaryFile1/27Sp_27Inds_20790G_Part1.tre')
tree<-drop.tip(tree, c('SpirodelaPolyrhiza'))

index = sort_data_by_treeTips(rownames(data_family_daf), tree)
data_family_daf=data_family_daf[index,]

```

Plot genetic diversity (pi4) against LHT

```

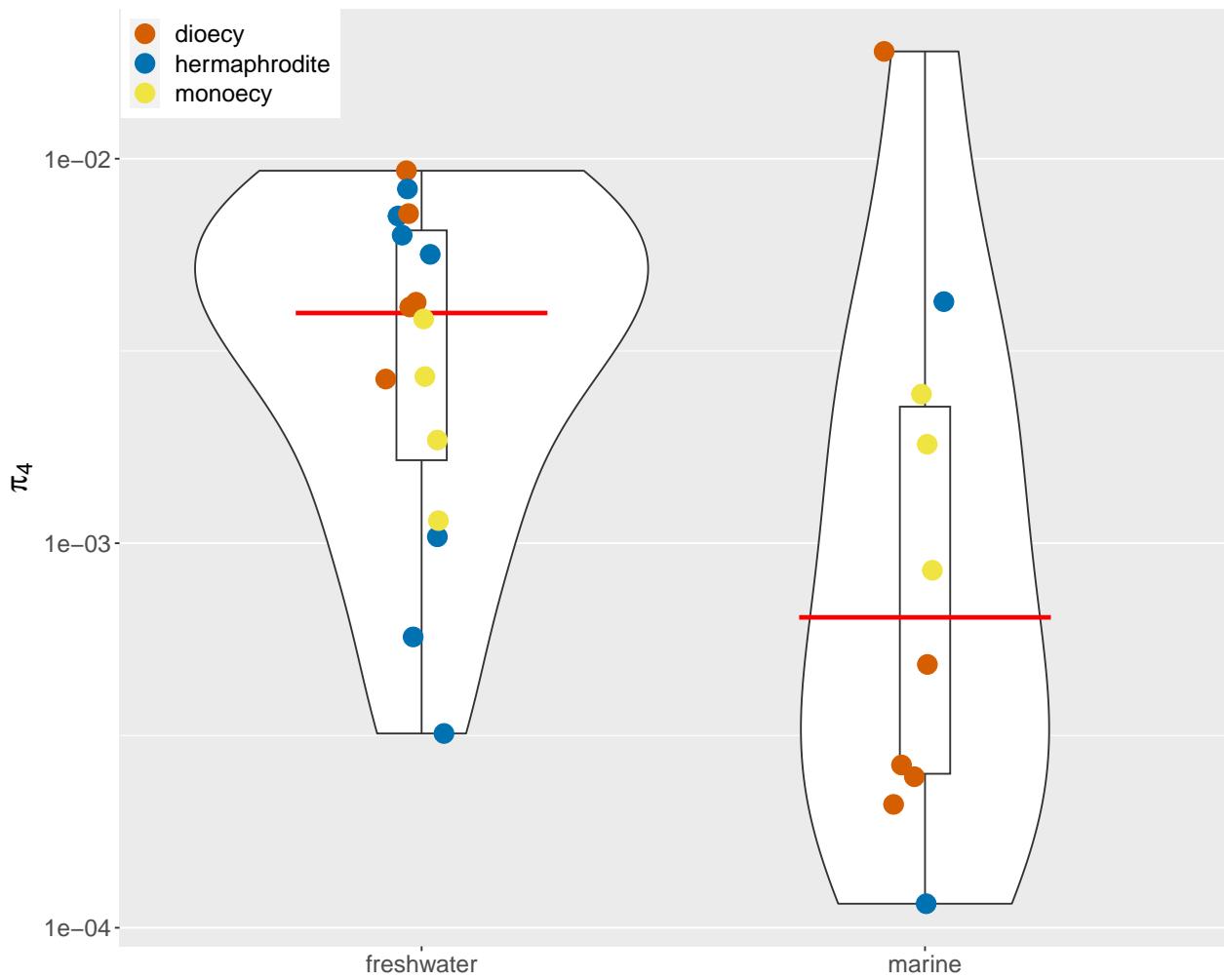
p<- ggplot(aes(x = habitat, y = pi4),data=data_family_daf) +
  geom_violin() +

```

```

geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar",
width = 0.5, colour = "red") +
geom_point(aes(fill = sex, color=sex), size = 5, shape = 21,
position = position_jitteredodge(dodge.width=0.1, jitter.width=0.1)) +
scale_y_log10() +
theme(text = element_text(size = 18), axis.title.x = element_blank(),
panel.grid.minor.x = element_blank(),
panel.grid.major.x = element_blank())+
ylab(expression(pi[4]))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy'="#F0E442"))+
scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy'="#F0E442"))+
theme(legend.position=c(0.1, 0.95), legend.title = element_blank())

```



```

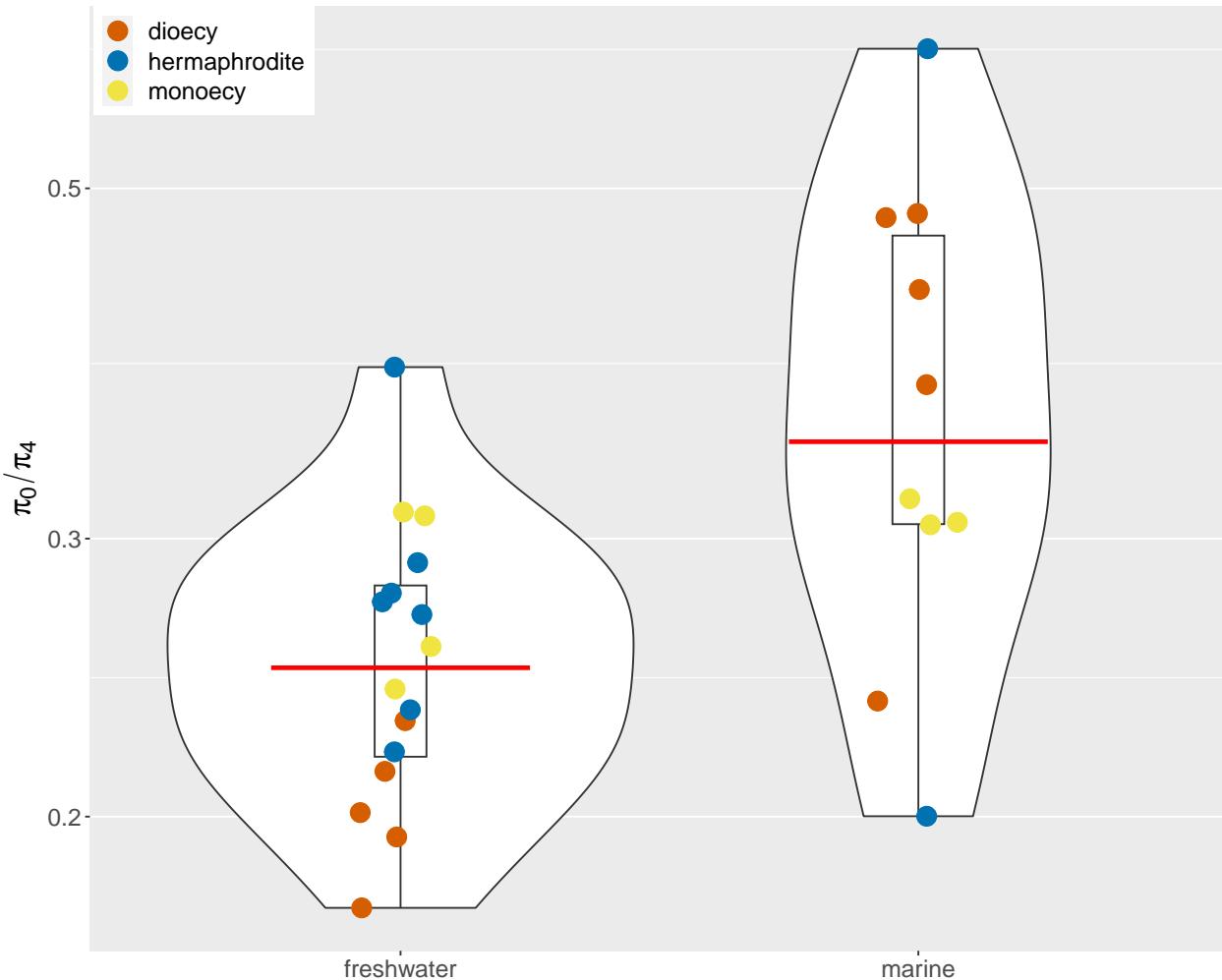
p<- ggplot(aes(x = habitat, y = pi0_pi4),data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar",
width = 0.5, colour = "red") +
  geom_point(aes(fill = sex, color=sex), size = 5, shape = 21,
position = position_jitteredodge(dodge.width=0.1, jitter.width=0.1)) +
  scale_y_log10() +
  theme(text = element_text(size = 18),axis.title.x = element_blank(),
  panel.grid.minor.x = element_blank(),

```

```

  panel.grid.major.x = element_blank())+
  ylab(expression(pi[0]/pi[4]))
p + scale_fill_manual(values=c('dioecy'="#D55E00", 'hermaphrodite'="#0072B2", 'monoecy'="#F0E442"))+
  scale_color_manual(values=c('dioecy'="#D55E00", 'hermaphrodite'="#0072B2", 'monoecy'="#F0E442"))+
  theme(legend.position=c(0.1, 0.95),legend.title = element_blank())

```



It seems there is drastic difference between dioecious species living in marine and fresh water

```

# test for significance
tapply(data_family_daf$pi4, list(data_family_daf$habitat, data_family_daf$sex), median)

##          dioecy   hermaphrodite   monoecy
## freshwater 0.004233658  0.005639544 0.002283538
## marine     0.000265000  0.002182229 0.001808461

m1=pglmm_compare(log10(pi4)~habitat*sex,phy=tree,data=data_family_daf,REML=F)

## as(<matrix>, "dgTMatrix") is deprecated since Matrix 1.5-0; do as(as(as(., "dMatrix"), "generalMatrix")
pglmm_compare(log10(pi4)~habitat,phy=tree,data=data_family_daf[data_family_daf$sex=='dioecy',],REML=F)

## Warning in pglmm_compare(log10(pi4) ~ habitat, phy = tree, data = data_family_daf[data_family_daf$se
## It appears that there are some species in phy are not contained in the rownames of data;
##           we will drop these species

```

```

## Linear mixed model fit by maximum likelihood
##
## Call:log10(pi4) ~ habitat
##
## logLik      AIC      BIC
## -8.035 24.070 20.701
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2        0.000  0.0000
## residual  0.292  0.5404
##
## Fixed effects:
##             Value Std.Error Zscore Pvalue
## (Intercept) -2.30125  0.24167 -9.5222 < 2e-16 ***
## habitatmarine -0.87865  0.34177 -2.5708 0.01015 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
tapply(data_family_daf$pi0_pi4, list(data_family_daf$habitat, data_family_daf$sex),median)

##          dioecy hermaphrodite monoecy
## freshwater 0.2011706    0.2735730 0.2832391
## marine     0.4314159    0.4065251 0.3072211
pglmm_compare(log10(pi0_pi4)~habitat,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:log10(pi0_pi4) ~ habitat
##
## logLik      AIC      BIC
## 20.32 -32.63 -32.18
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2        0.00000  0.0000
## residual  0.01227  0.1108
##
## Fixed effects:
##             Value Std.Error Zscore Pvalue
## (Intercept) -0.601895  0.027692 -21.7357 < 2.2e-16 ***
## habitatmarine 0.153059  0.044651   3.4279 0.0006083 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
pglmm_compare(log10(pi0_pi4)~habitat,phy=tree,data=data_family_daf[data_family_daf$sex=='dioecy',],REML=F)

## Warning in pglmm_compare(log10(pi0_pi4) ~ habitat, phy = tree, data = data_family_daf[data_family_daf$sex=='dioecy', ]):
## It appears that there are some species in phy are not contained in the rownames of data;
## we will drop these species

## Linear mixed model fit by maximum likelihood
##
## Call:log10(pi0_pi4) ~ habitat
##
## logLik      AIC      BIC

```

```

##  10.37 -12.74 -16.11
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.000000 0.00000
## residual 0.007358 0.08578
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) -0.694715  0.038362 -18.1093 < 2.2e-16 ***
## habitatmarine 0.284168  0.054253   5.2379 1.624e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
pglmm_compare(log10(pi0_pi4)~habitat*sex,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:log10(pi0_pi4) ~ habitat * sex
##
## logLik     AIC     BIC
##  24.11 -32.22 -31.31
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.000000 0.00000
## residual 0.009165 0.09573
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) -0.694715  0.042814 -16.2265 < 2.2e-16 ***
## habitatmarine 0.284168  0.060548   4.6933 2.688e-06 ***
## sexhermaphrodite 0.132872  0.056056   2.3703 0.01777 *
## sexmonoecy    0.138753  0.064220   2.1606 0.03073 *
## habitatmarine:sexhermaphrodite -0.177983  0.097764  -1.8205 0.06868 .
## habitatmarine:sexmonoecy      -0.236310  0.094933  -2.4892 0.01280 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# phylog effect
x= data_family_daf$pi4
names(x)=rownames(data_family_daf)
m1 = phylosig(tree, log10(x),test=T)
print(m1)

##
## Phylogenetic signal K : 0.489139
## P-value (based on 1000 randomizations) : 0.035

```

## Chapter 2. DFE

```

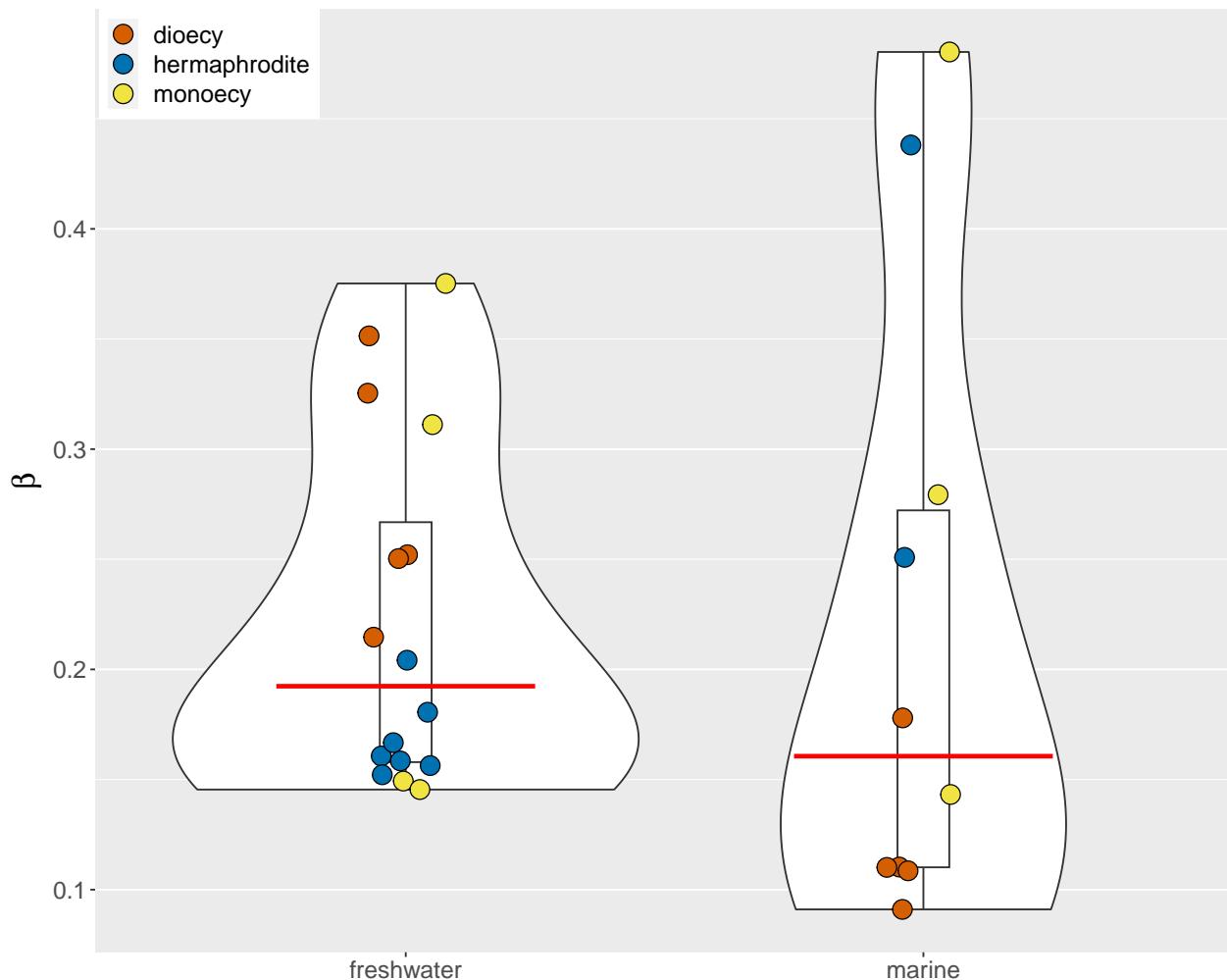
p<- ggplot(aes(x = habitat, y = b),data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar", width = 0.5,colour = "red")

```

```

geom_point(aes(fill = sex), size = 5, shape = 21, position = position_jitterdodge(dodge.width=0.1, ji
theme(text = element_text(size = 18),
      axis.title.x = element_blank(),
      panel.grid.minor.x = element_blank(),
      panel.grid.major.x = element_blank(),
    ) +
  ylab(expression(beta))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F0E442"))+
  scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F0E442"))
  theme(legend.position=c(0.1, 0.95), legend.title = element_blank())

```



```

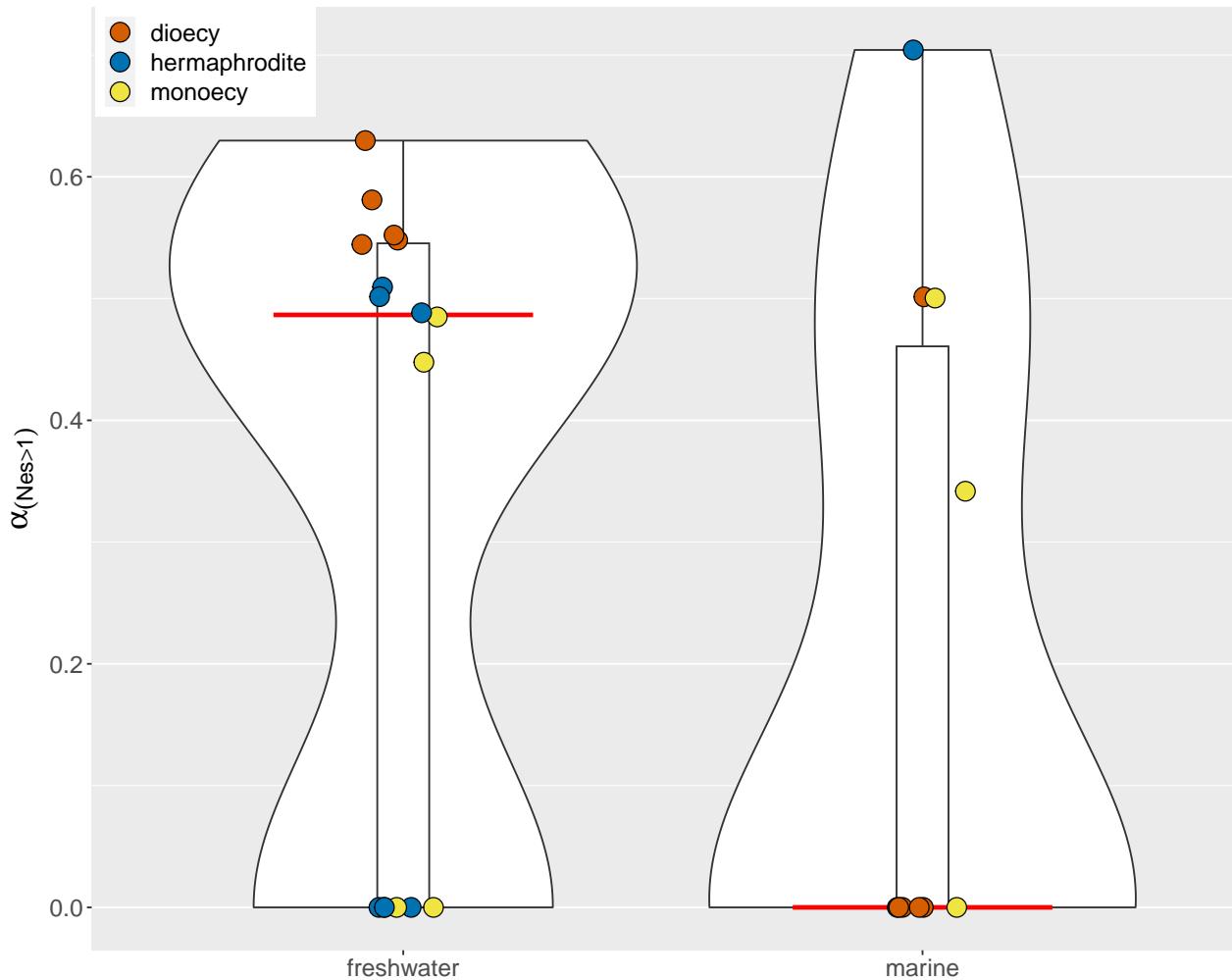
p<-ggplot(aes(x = habitat, y =`alpha > 1`),data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar", width = 0.5,co
  geom_point(aes(fill = sex), size = 5, shape = 21, position = position_jitterdodge(dodge.width=0.1, ji
  theme(text = element_text(size = 18),
        axis.title.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.x = element_blank(),
      ) +
  ylab(expression(alpha[(Nes>1)])))
  p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F0E442"))

```

```

    scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy'="#F08080")
    theme(legend.position=c(0.1, 0.95),legend.title = element_blank())

```



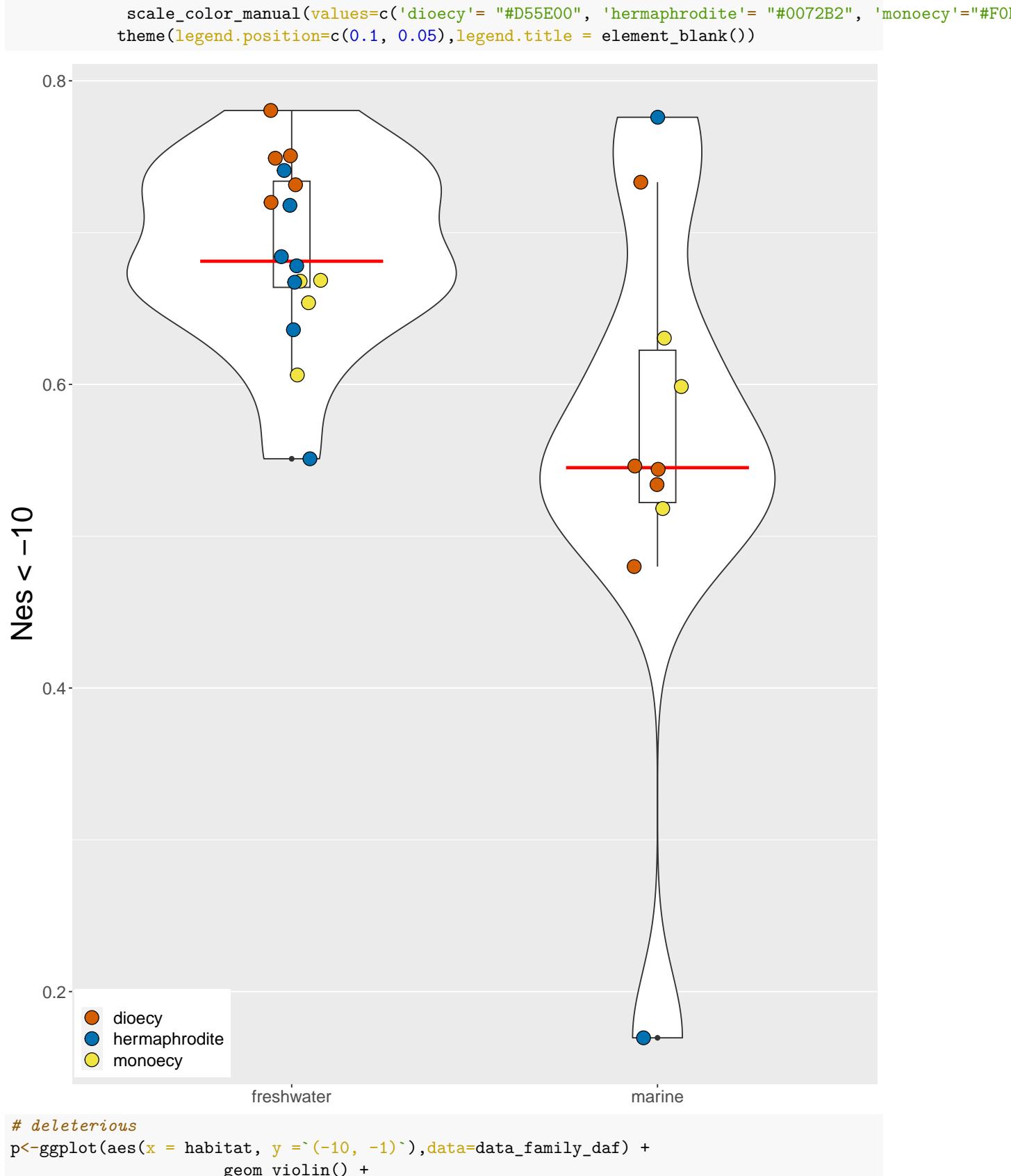
```

# strong deleterious mutations
data_family_daf$str_del= data_family_daf$`<-100` + data_family_daf$`(-100, -10)`

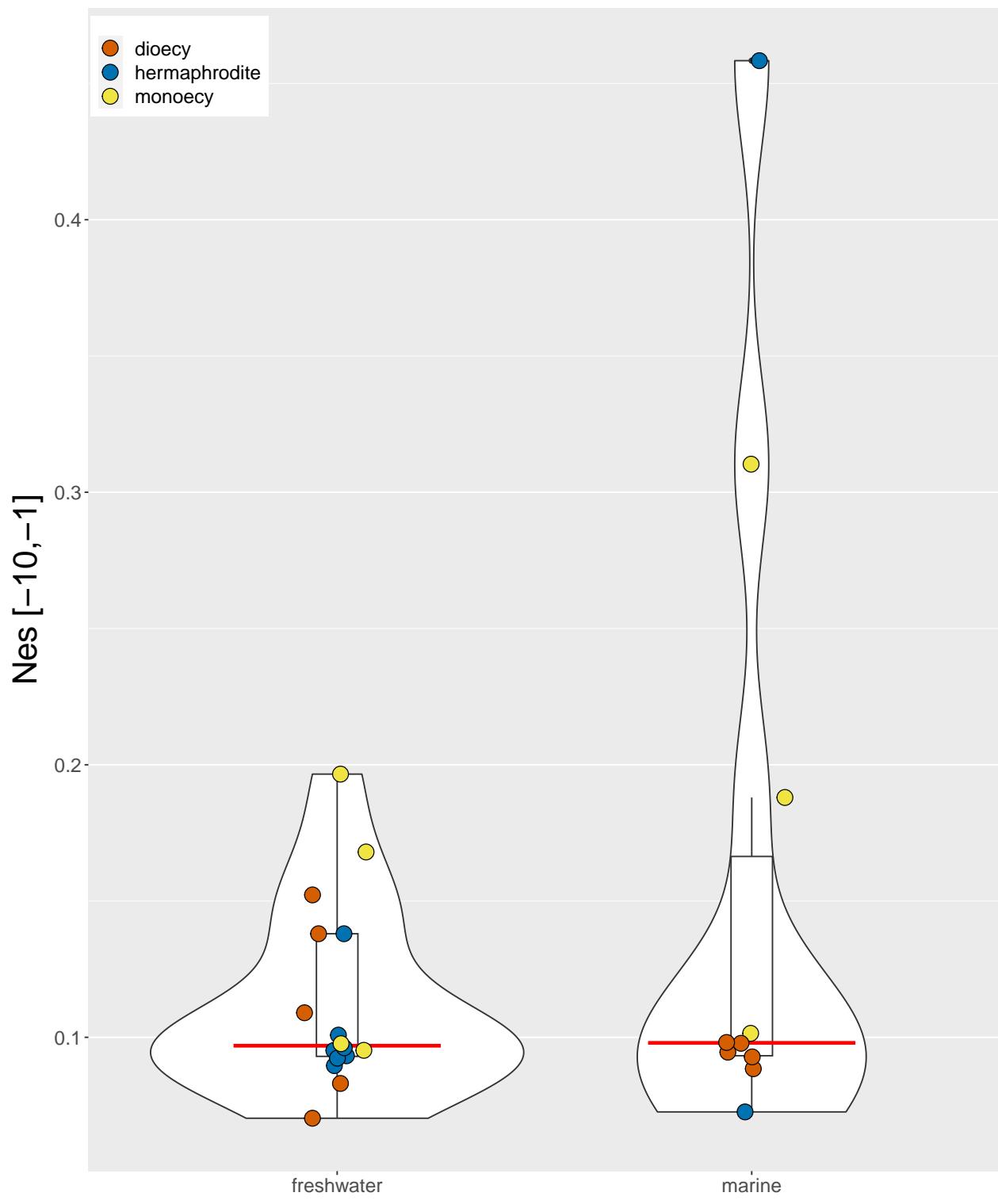
# strong beneficial
data_family_daf$str_benf= data_family_daf$`(1, 10)` + data_family_daf$`10<` 

# strongly deleterious
p<-ggplot(aes(x = habitat, y =str_del),data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar", width = 0.5,color="red")
  geom_point(aes(fill = sex), size = 5, shape = 21, position = position_jitterdodge(dodge.width=0.5))
  theme(text = element_text(size = 18),
        axis.title.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.x = element_blank(),
        axis.title=element_text(size=24,face="bold"))
  ) +
  ylab(expression('Nes < -10'))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy'="#F08080"))

```

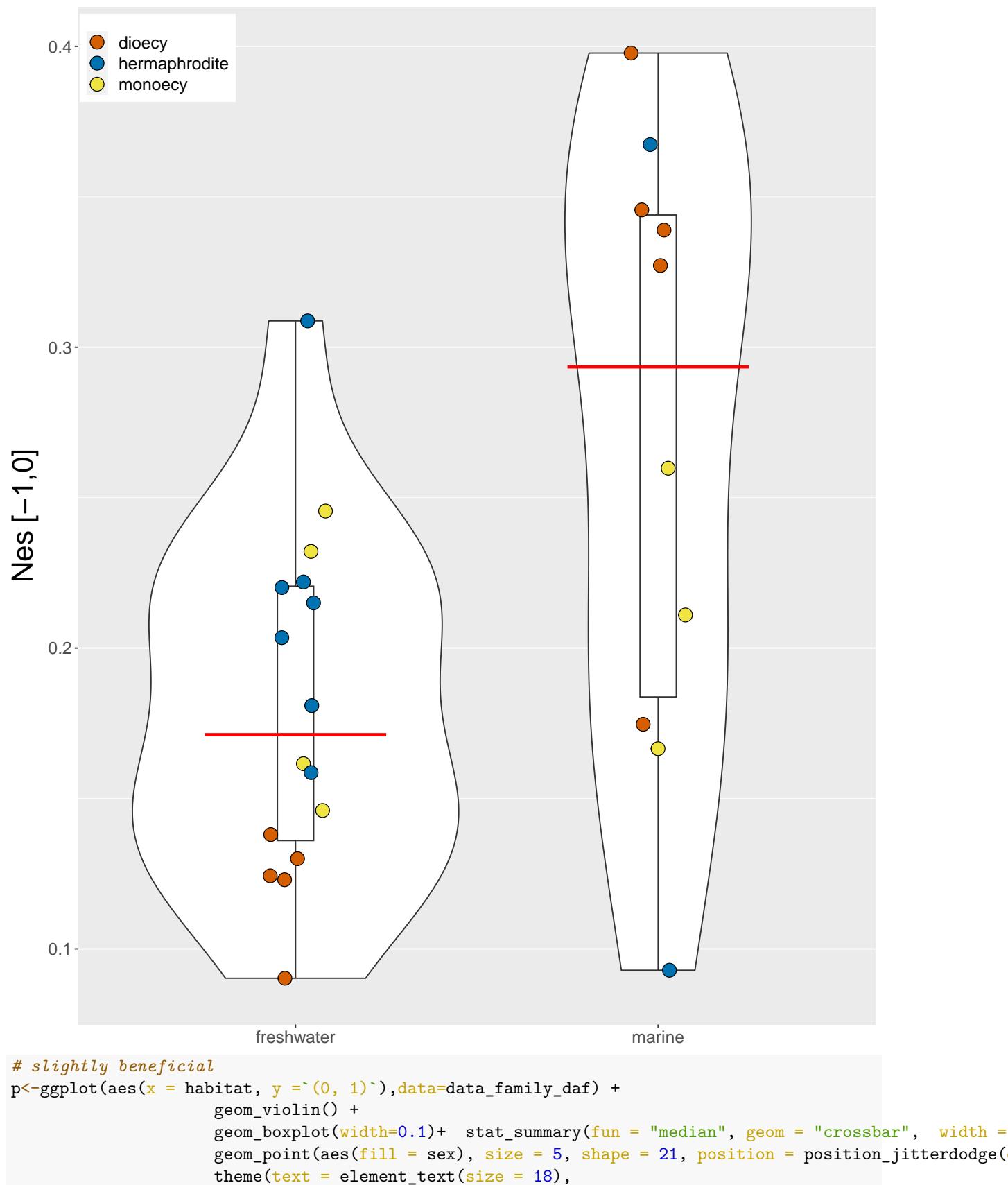


```
geom_boxplot(width=0.1)+  stat_summary(fun = "median", geom = "crossbar", width =
geom_point(aes(fill = sex), size = 5, shape = 21, position = position_jitterdodge
theme(text = element_text(size = 18),
      axis.title.x = element_blank(),
      panel.grid.minor.x = element_blank(),
      panel.grid.major.x = element_blank(),
      axis.title=element_text(size=24,face="bold")
) +
ylab(expression('Nes [-10,-1]'))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'mono
scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'mono
theme(legend.position=c(0.1, 0.95),legend.title = element_blank())
```

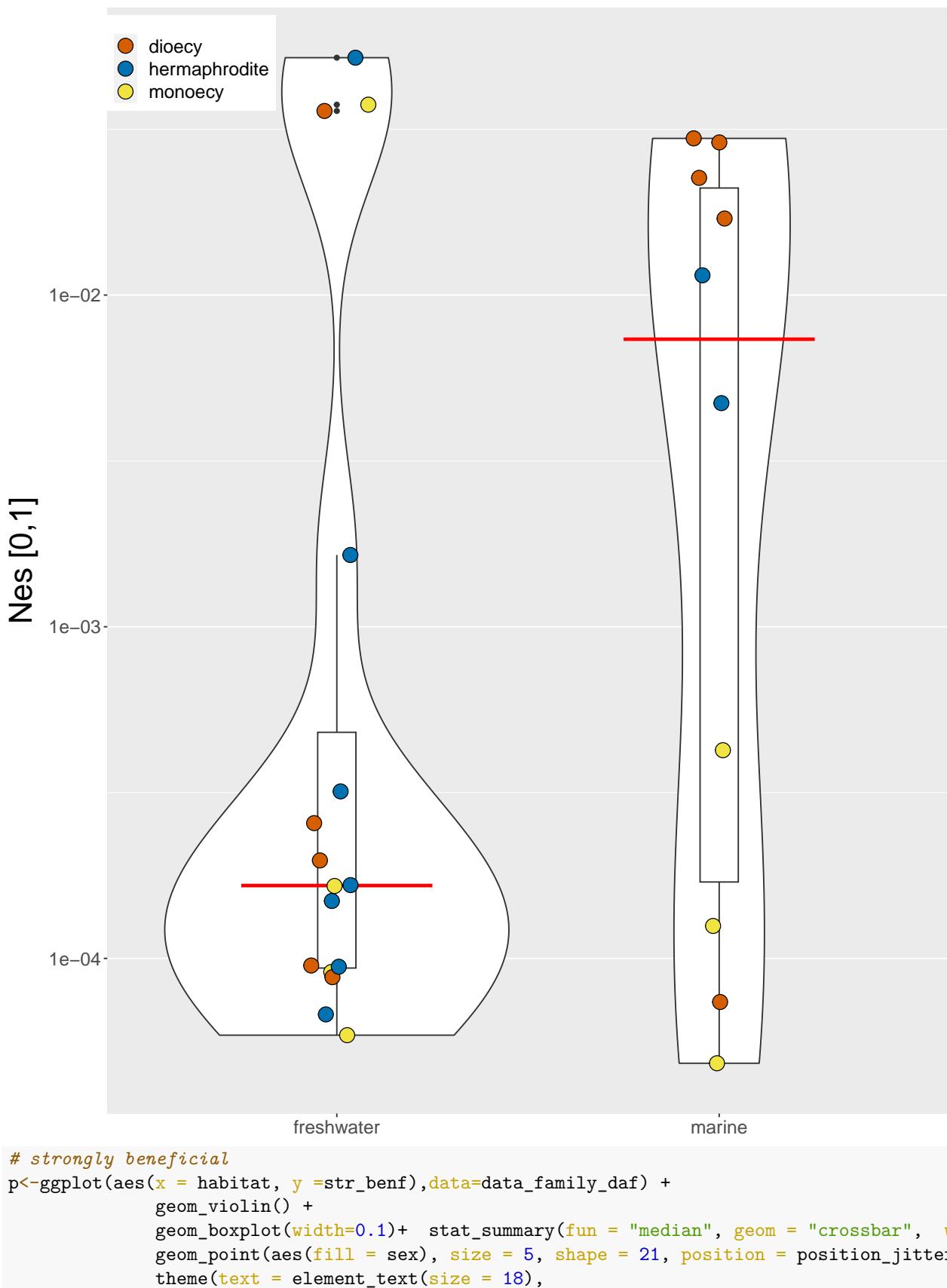


```
# neutral or nearly neutral
p<-ggplot(aes(x = habitat, y = `(-1,0)`), data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar", width =
  geom_point(aes(fill = sex), size = 5, shape = 21, position = position_jitterdodge(
  theme(text = element_text(size = 18),
```

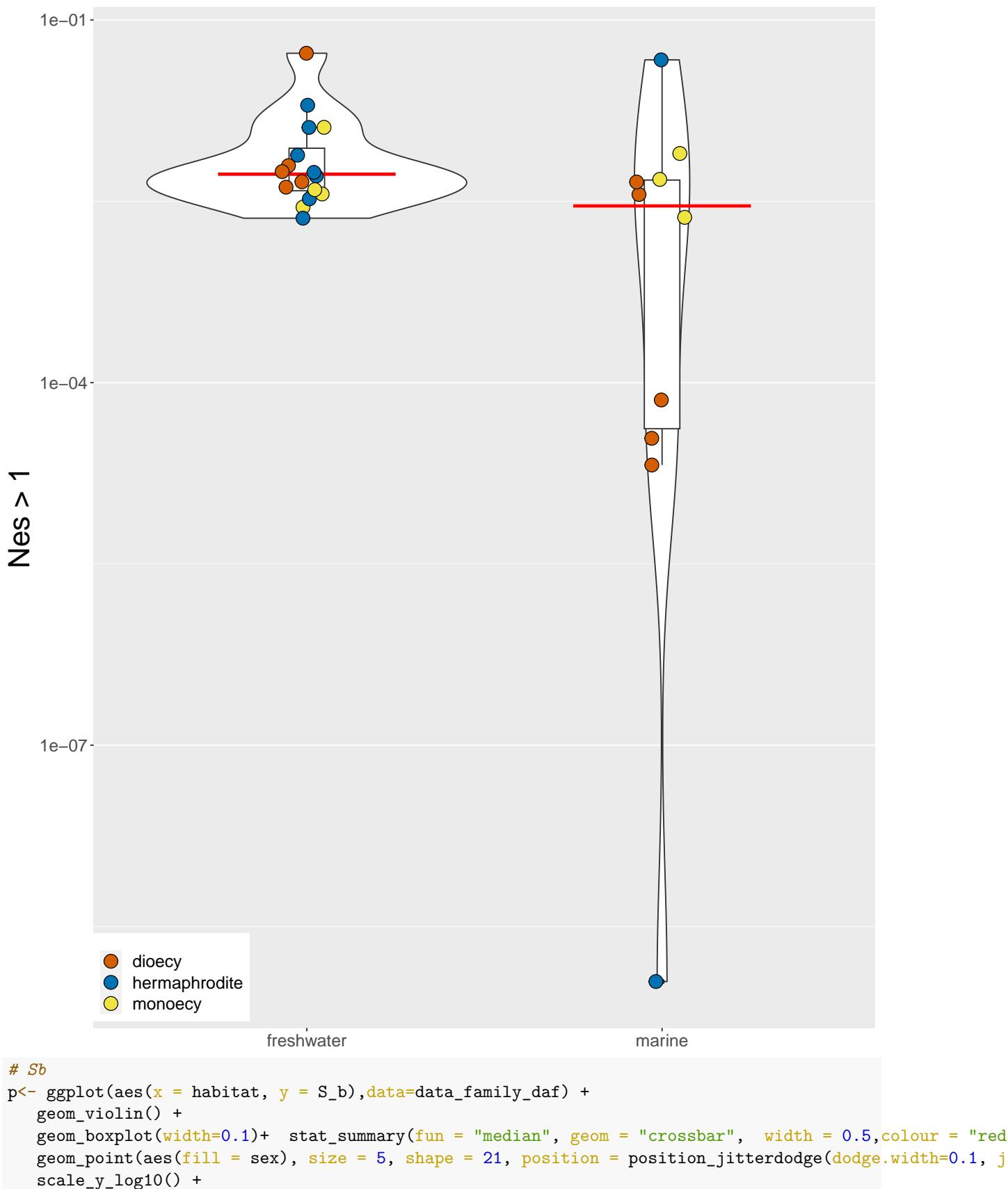
```
axis.title.x = element_blank(),
panel.grid.minor.x = element_blank(),
panel.grid.major.x = element_blank(),
axis.title=element_text(size=24,face="bold")
) +
ylab(expression('Nes [-1,0]'))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F0E442"))+
  scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'mon
theme(legend.position=c(0.1, 0.95),legend.title = element_blank())
```



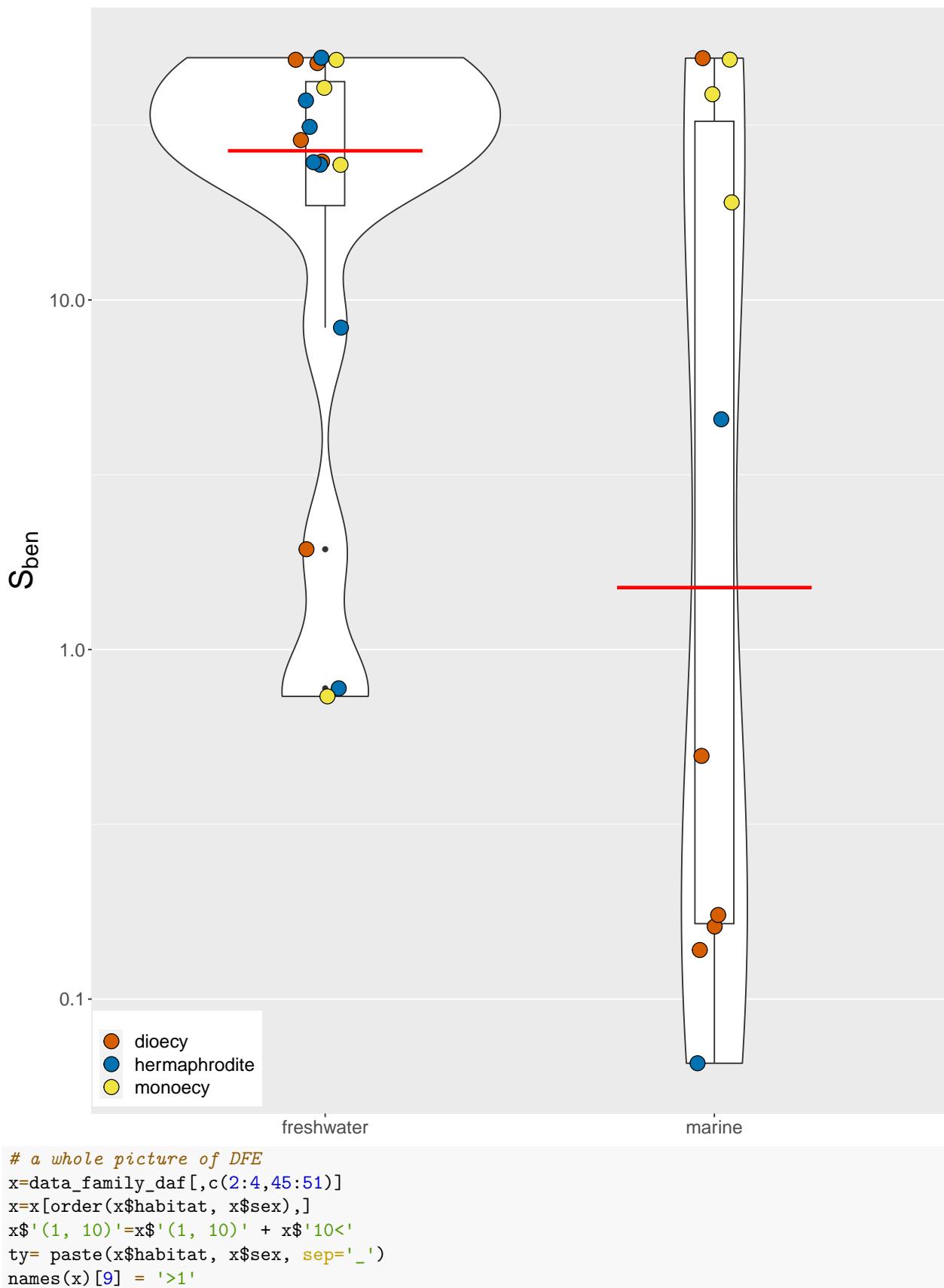
```
axis.title.x = element_blank(),
panel.grid.minor.x = element_blank(),
panel.grid.major.x = element_blank(),
axis.title=element_text(size=24,face="bold")
) +    scale_y_log10() +
ylab(expression('Nes [0,1]'))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F0E442"))+
  scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'mon
theme(legend.position=c(0.1, 0.95),legend.title = element_blank())
```



```
axis.title.x = element_blank(),
panel.grid.minor.x = element_blank(),
panel.grid.major.x = element_blank(),
axis.title=element_text(size=24,face="bold")
) + scale_y_log10() +
ylab(expression('Nes > 1'))
p + scale_fill_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F08080"))
scale_color_manual(values=c('dioecy' = "#D55E00", 'hermaphrodite' = "#0072B2", 'monoecy' = "#F08080"))
theme(legend.position=c(0.1, 0.05),legend.title = element_blank())
```



```
theme(text = element_text(size = 18),
      axis.title.x = element_blank(),
      panel.grid.minor.x = element_blank(),
      panel.grid.major.x = element_blank(),
      axis.title=element_text(size=24,face="bold")) +
  ylab(expression(S[ben]))  
p + scale_fill_manual(values=c('dioecy'= "#D55E00", 'hermaphrodite'= "#0072B2", 'monoecy'="#F0E442"))+  
  scale_color_manual(values=c('dioecy'= "#D55E00", 'hermaphroditic'= "#0072B2", 'monoecy'="#F0E442"))+  
  theme(legend.position=c(0.1, 0.05),legend.title = element_blank())
```



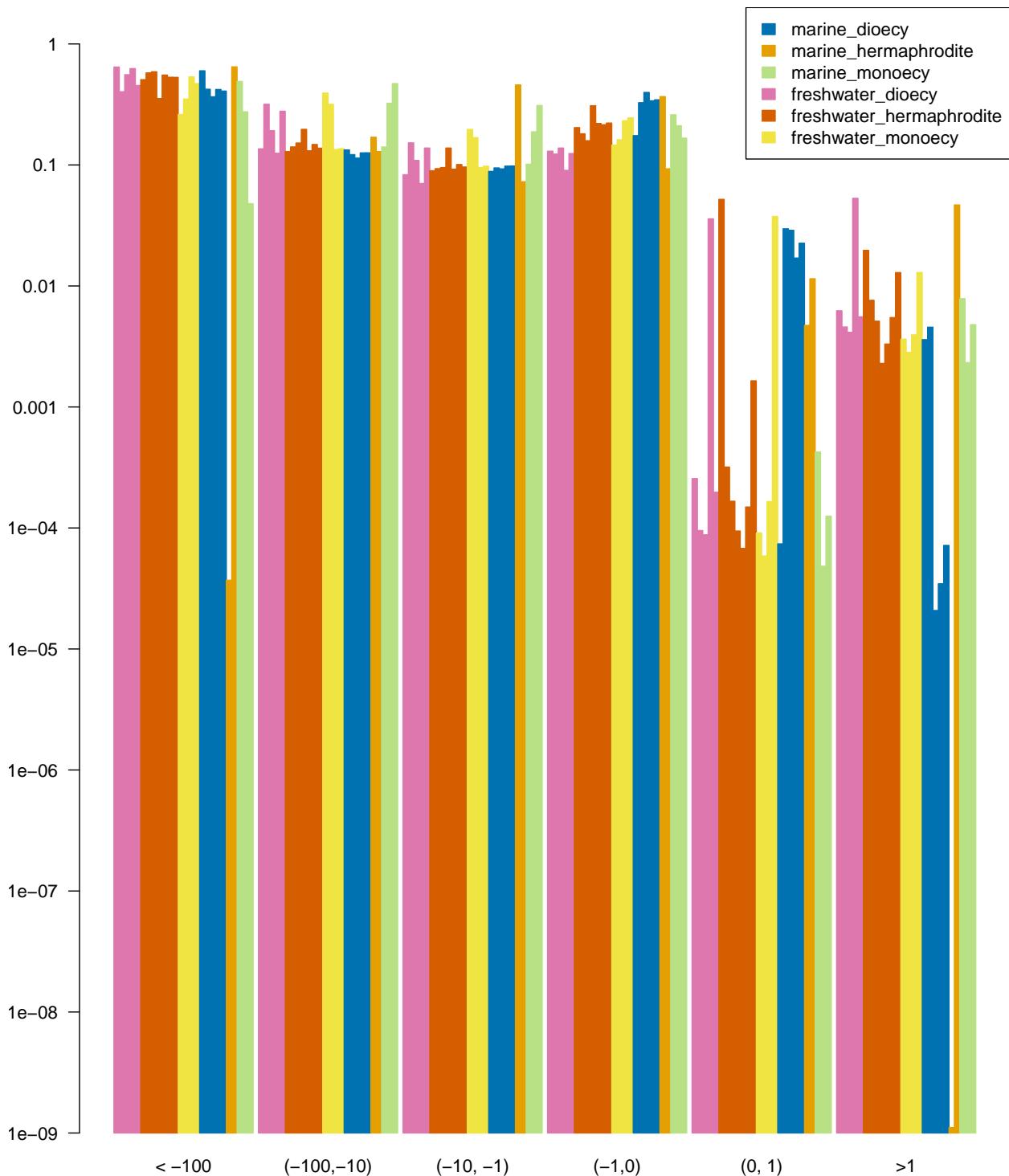
```

names(x)[4] = '< -100'

col.vector = c('#0072b2', '#e69f00', '#b8e186', '#de77ae', '#d55e00', '#f0e442')
col_v = ifelse(ty=='marine_dioecy', col.vector[1],
               ifelse(ty=='marine_hermaphrodite', col.vector[2],
               ifelse(ty=='marine_monoecy', col.vector[3],
               ifelse(ty=='freshwater_dioecy', col.vector[4],
               ifelse(ty=='freshwater_hermaphrodite', col.vector[5],col.vector[6])))))

# barplot(as.matrix(x[,4:9]),beside=T, col=col_v, log='y')
par(mar=c(5,5,3,2))
barplot(as.matrix(x[,4:9]),beside=T, col=col_v, log='y',border=col_v, yaxt = 'n', ylim=c(1e-9,2))
axis(2, at=c(1e-9,1e-8,1e-7,1e-6,1e-5,1e-4,1e-3,1e-2,1e-1,1),lab=c(1e-9,1e-8,1e-7,1e-6,1e-5,1e-4,1e-3,1e-2,1e-1,1))
legend('topright', legend=c('marine_dioecy', 'marine_hermaphrodite', 'marine_monoecy', 'freshwater_dioecy'))

```



```
##  
tapply(data_family_daf$b, list(data_family_daf$habitat, data_family_daf$sex), median)
```

```
##          dioecy hermaphrodite monoecy  
## freshwater 0.2520644      0.1606814 0.2302210  
## marine    0.1101105      0.3444741 0.2793139
```

```

pglmm_compare(`b` ~ habitat, phy=tree, data=data_family_daf[data_family_daf$sex=='dioecy',], REML=F)

## Warning in pglmm_compare(b ~ habitat, phy = tree, data = data_family_daf[data_family_daf$sex == :
## It appears that there are some species in phy are not contained in the rownames of data;
## we will drop these species

## Linear mixed model fit by maximum likelihood
##
## Call:b ~ habitat
##
## logLik      AIC      BIC
## 17.52 -27.04 -30.41
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.000000 0.00000
## residual 0.001761 0.04197
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) 0.278766 0.018768 14.8532 < 2.2e-16 ***
## habitatmarine -0.159176 0.026542 -5.9971 2.009e-09 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

## Chapter 3. Effects of distribution range, Nc and historical Ne

We further investigated the correlation of genetic diversity and census size

Estimate historical Ne

```

## read stairway plot2 results
file_all_sites = list.files(path='~/Desktop/ZJU/seagrass/1-manuscripts/202511/supplementaryFile1/staiway')

n=length(file_all_sites)
res.allSites = list()
x.max=NULL
y.min=NULL
y.max=NULL
for(i in 1:(n+1)){
  res.allSites[i]=NULL
}
for(i in 1:n){
  demo=read.table(file_all_sites[i], header=T)
  id = strsplit(basename(file_all_sites[i]), '_')[[1]][1]
  res.allSites[[i]]$id = id
  res.allSites[[i]]$habitat = data_family_daf$habitat[data_family_daf$Species==id]
  res.allSites[[i]]$sex = data_family_daf$sex[data_family_daf$Species==id]
  res.allSites[[i]]$Ne=demo$Ne_median
  res.allSites[[i]]$year=demo$year
  x.max=c(x.max, max(demo$year))
  y.min=c(y.min, min(demo$Ne_median))
  y.max=c(y.max, max(demo$Ne_median))
}

```

```

year.max = max(x.max)
Ne.min=min(y.min)
Ne.max=max(y.max)

```

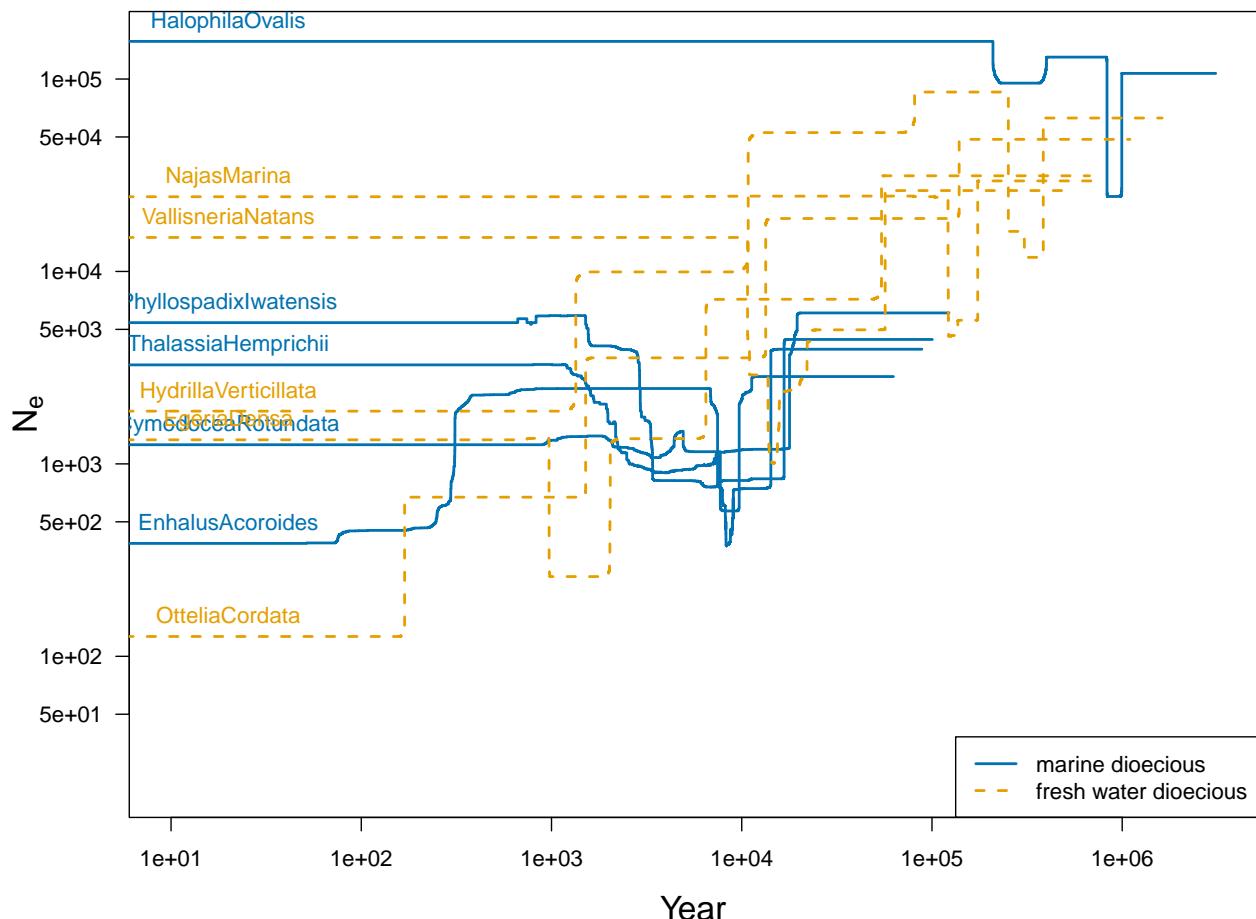
Plot stairway plot results

```

par(mar=c(6,6,4,2))
plot(c(10, year.max), c(Ne.min, Ne.max), type='n', ylab=expression(N[e]), xlab='Year', log='xy', las=1,
for(i in 1:n){
  if(res.allSites[[i]]$habitat=='marine' & res.allSites[[i]]$sex=='dioecy'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#0072B2', lwd=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3, off=.5
  }
}

for(i in 1:n){
  if(res.allSites[[i]]$habitat=='freshwater' & res.allSites[[i]]$sex=='dioecy'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#E69F00',lwd=2, lty=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3, off=.5
  }
}
legend('bottomright', legend=c('marine dioecious', 'fresh water dioecious'), lty=c(1,2),lwd=2, col=c('#0072B2', '#E69F00'))

```



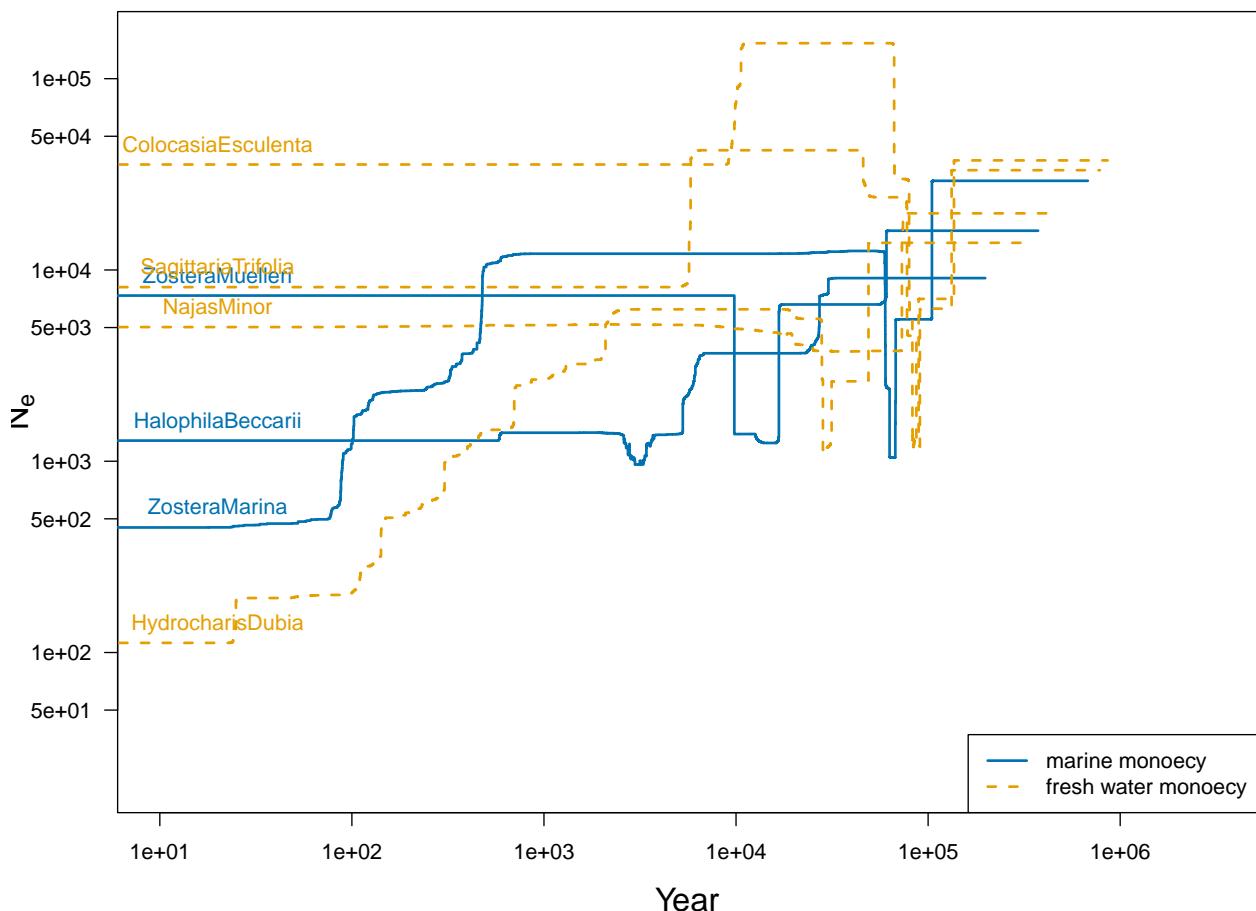
Historical Ne for marine and fresh water dioecious species. We observed significantly lower historical Ne for marine dioecious species, though they were also relatively stable.

```

plot(c(10, year.max), c(Ne.min, Ne.max), type='n', ylab=expression(N[e]), xlab='Year', log='xy', las=1,
for(i in 1:n){
  if(res.allSites[[i]]$habitat=='marine' & res.allSites[[i]]$sex=='monoecy'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#0072B2', lwd=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3,off=.5)
  }
}

for(i in 1:n){
  if(res.allSites[[i]]$habitat=='freshwater' & res.allSites[[i]]$sex=='monoecy'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#E69F00',lwd=2, lty=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3,off=.5)
  }
}
legend('bottomright', legend=c('marine monoecy', 'fresh water monoecy'), lty=c(1,2),lwd=2, col=c('#0072B2', '#E69F00'))

```



```

## 
plot(c(10, year.max), c(Ne.min, Ne.max), type='n', ylab=expression(N[e]), xlab='Year', log='xy', las=1,
for(i in 1:n){
  if(res.allSites[[i]]$habitat=='marine' & res.allSites[[i]]$sex=='hermaphrodite'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#0072B2', lwd=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3,off=.5)
  }
}

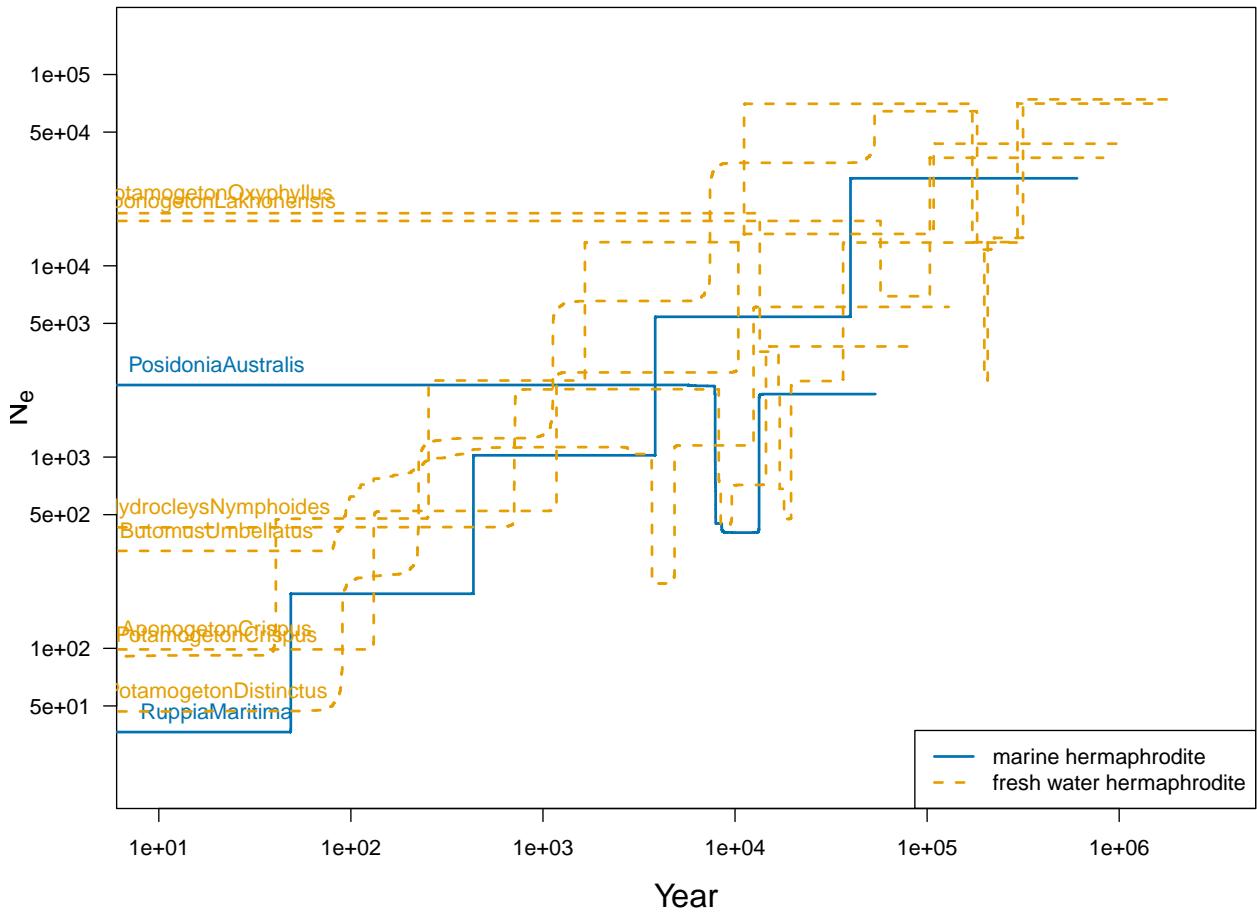
```

```

}

for(i in 1:n){
  if(res.allSites[[i]]$habitat=='freshwater' & res.allSites[[i]]$sex=='hermaphrodite'){
    lines(res.allSites[[i]]$year, res.allSites[[i]]$Ne,col= '#E69F00',lwd=2, lty=2)
    text(20, res.allSites[[i]]$Ne[res.allSites[[i]]$year>20][1], res.allSites[[i]]$id, pos=3,off=1)
  }
}
legend('bottomright', legend=c('marine hermaphrodite', 'fresh water hermaphrodite'), lty=c(1,2),lwd=2, col=c('blue','orange'))

```



For monoecious and hermaphroditic species. *Posidonia australis* also had a lower Ne curve which corresponds to its small census size. We further estimated a harmonic mean of historical Ne, which should be positively correlated to pi4.

```

##### calculate Ne
ids = NULL
Ne_harmonic = NULL
for(i in 1:n){
  index = res.allSites[[i]]$year > 1e2 & res.allSites[[i]]$year < 1e6
  ids = c(ids, res.allSites[[i]]$id)
  m = sum(index)
  Ne = res.allSites[[i]]$Ne[index]
  t = res.allSites[[i]]$year[index]
  Ne_v = Ne[1] # to restore different Ne
  t_v = t[1]
  for(j in 2:m){
    Ne_v = Ne_v * t[j]/t_v
    t_v = t_v * t[j]
    Ne_harmonic = c(Ne_harmonic, Ne_v)
  }
}

```

```

if(Ne[j] / tail(Ne_v,1) >= 2 | Ne[j] / tail(Ne_v,1) <= 0.5){
  Ne_v = c(Ne_v, Ne[j])
  t_v = c(t_v, t[j])
}
}
Ne_v = c(Ne_v, Ne[m])
t_v = c(t_v, t[m])
l = length(Ne_v)
Ne_harmonic_inverse_sum = 0
for(k in 1:(l-1)){
  delta_t = t_v[k+1] - t_v[k]
  Ne_harmonic_inverse_sum = Ne_harmonic_inverse_sum + delta_t / Ne_v[k]
}
t_total = t_v[l] - t_v[1]
Ne_harmonic = c(Ne_harmonic, t_total / Ne_harmonic_inverse_sum)
# print(paste(id, round(Ne_harmonic)))
}
Ne = data.frame(ids, Ne_harmonic)

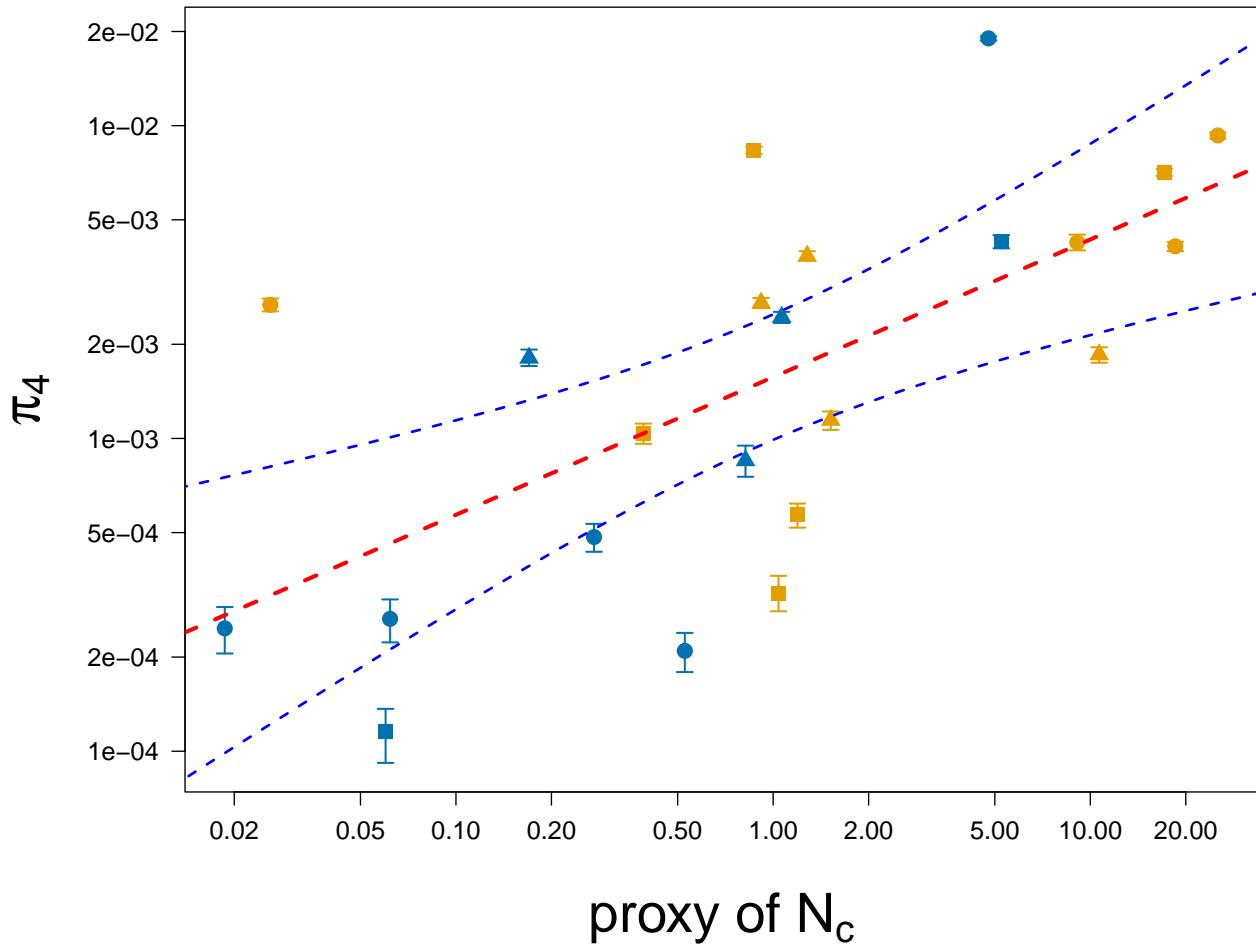
data_family_daf$Ne_harmonic = NA
for(i in 1:nrow(data_family_daf)){
  id=rrownames(data_family_daf)[i]
  if(sum(Ne$ids==id)){
    data_family_daf$Ne_harmonic[i] = Ne$Ne_harmonic[Ne$ids==id]
  }
}

#
pi4_lw = as.numeric(unlist(lapply(strsplit(data_family_daf$'pi4_95%', ',', ''), function(a){a[1]})))
pi4_hi = as.numeric(unlist(lapply(strsplit(data_family_daf$'pi4_95%', ',', ''), function(a){a[2]})))

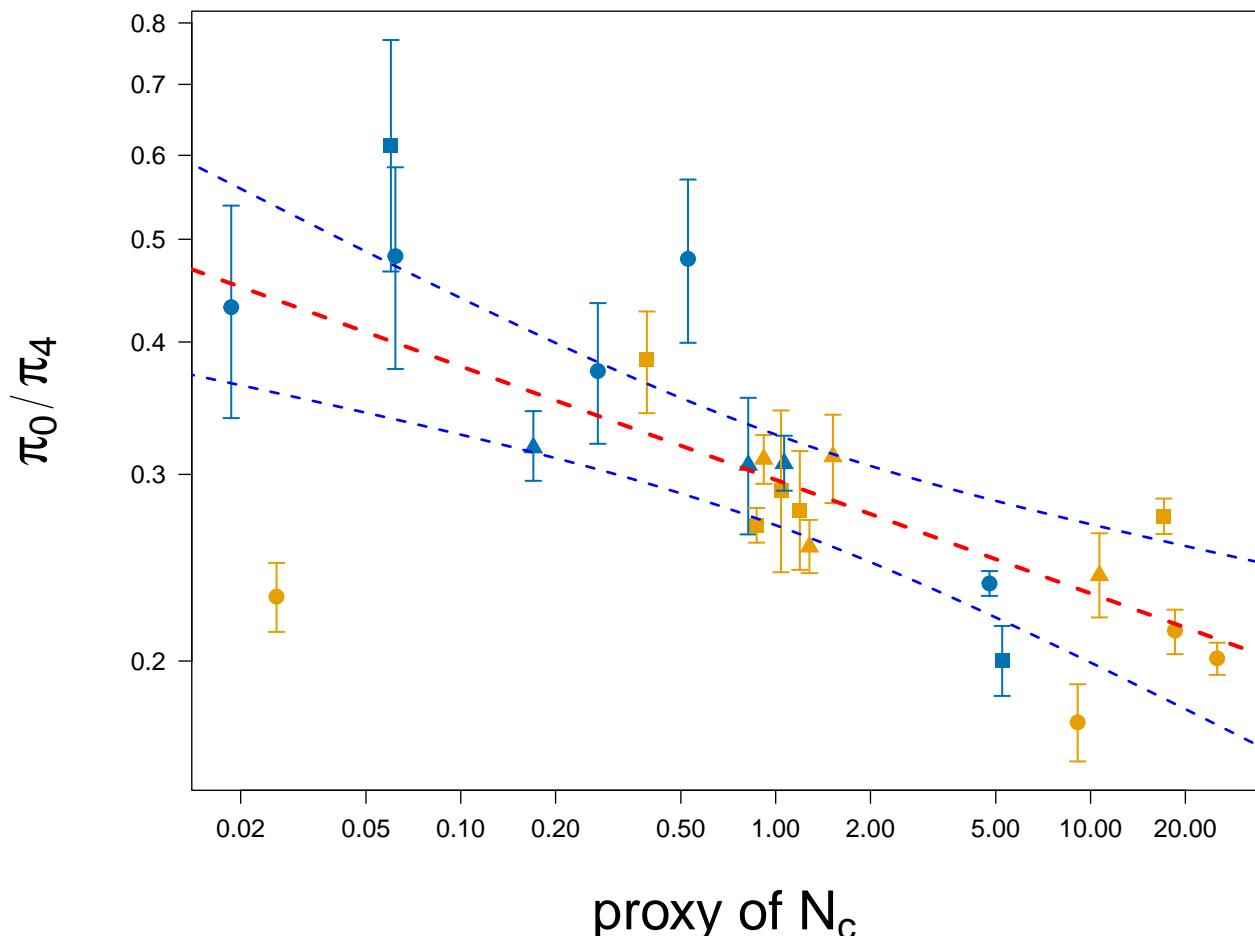
pi0_pi4_lw = as.numeric(unlist(lapply(strsplit(data_family_daf$'pi0_pi4_95%', ',', ''), function(a){a[1]})))
pi0_pi4_hi = as.numeric(unlist(lapply(strsplit(data_family_daf$'pi0_pi4_95%', ',', ''), function(a){a[2]})))

par(mar=c(7,7,3,2), mgp=c(5,1,0),cex.lab=2.5,cex.axis=1.2, font=2, las=1)
plot(data_family_daf$census, data_family_daf$pi4, log='xy', xlab=expression(paste('proxy of ', N[c])), ylab='Ne_harmonic',
      pch=ifelse(data_family_daf$sex=='dioecy', 19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
      col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'),cex=1.5,ylim=c(min(pi4_lw),max(pi4_hi)))
# legend('topleft',legend=c('monoecy', 'dioecy', 'hermaphrodite', 'marine', 'freshwater'),
#        pch=c(17,19,15, 19, 19), col=c(1,1,1,'#0072B2','#E69F00'),cex=1.5)
plot_ci(data_family_daf$census, pi4_lw, pi4_hi, col=ifelse(data_family_daf$habitat=='marine', '#0072B2',
# text(data_family_daf$census, data_family_daf$pi4, rounames(data_family_daf), pos=3, offset=.5, cex=.4)
model<-lm(log10(data_family_daf$pi4)~log10(data_family_daf$census))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$census), model, log='xy',lwd=2,lty=2,col='blue')

```



```
par(mar=c(7,7,3,2), mgp=c(5,1,0), cex.lab=2.5, cex.axis=1.2, font=2, las=1)
plot(data_family_daf$census, data_family_daf$pi0_pi4, log='xy', xlab=expression(paste('proxy of ', N[c])),
      pch=ifelse(data_family_daf$sex=='dioecy', 19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
      col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'), cex=1.5, ylim=c(min(pi0_pi4_lw),max(pi0_pi4_hi)))
plot_ci(data_family_daf$census, pi0_pi4_lw, pi0_pi4_hi, col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'))
# text(data_family_daf$census, data_family_daf$pi0_pi4, rownames(data_family_daf), pos=3, offset=.5, cex=1.5)
model<-lm(log10(data_family_daf$pi0_pi4)~log10(data_family_daf$census))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$census), model, log='xy', lwd=2, lty=2, col='blue')
```



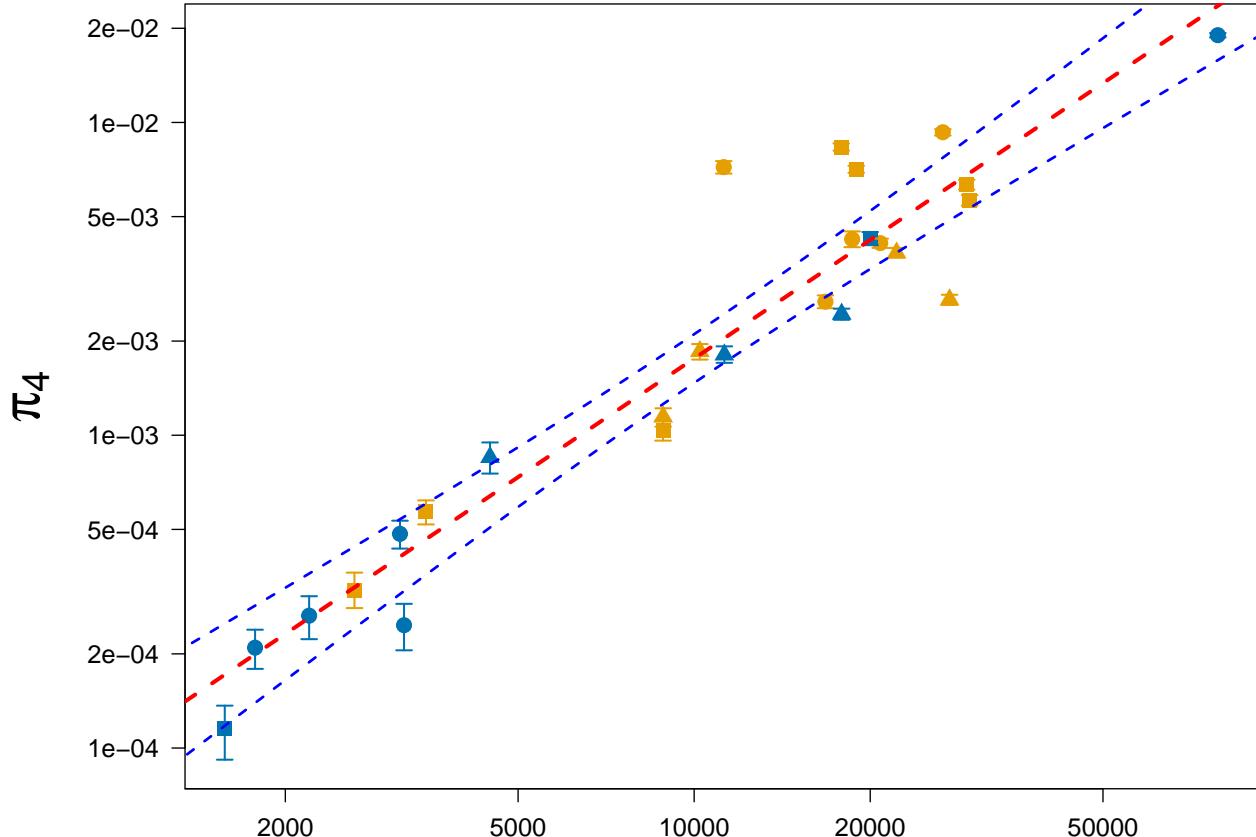
```
summary(model)

##
## Call:
## lm(formula = log10(data_family_daf$pi0_pi4) ~ log10(data_family_daf$census))
##
## Residuals:
##      Min        1Q        Median       3Q        Max 
## -0.279944 -0.037220 -0.006477  0.041816  0.184568 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -0.52803   0.02057 -25.669 < 2e-16 ***
## log10(data_family_daf$census) -0.10711   0.02343 -4.571 0.000166 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09865 on 21 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.4987, Adjusted R-squared:  0.4748 
## F-statistic: 20.89 on 1 and 21 DF,  p-value: 0.0001661
## 
## 
## par(mar=c(7,7,3,2), mgp=c(5,1,0),cex.lab=2.5,cex.axis=1.2, font=2, las=1)
plot(data_family_daf$Ne_harmonic, data_family_daf$pi4,log='xy', xlab=expression(N[e]),ylab=expression(p)
```

```

pch=ifelse(data_family_daf$sex=='dioecy',19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'),cex=1.5, ylim=c(min(pi4_lw),max(pi4_hi)),
# text(data_family_daf$Ne_harmonic, data_family_daf$pi4, rownames(data_family_daf), pos=3, offset=.5, cex=1.5),
plot_ci(data_family_daf$Ne_harmonic, pi4_lw, pi4_hi, col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00')),
model<-lm(log10(data_family_daf$pi4)~log10(data_family_daf$Ne_harmonic))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$Ne_harmonic), model, log='xy',lwd=2,lty=2,col='blue')

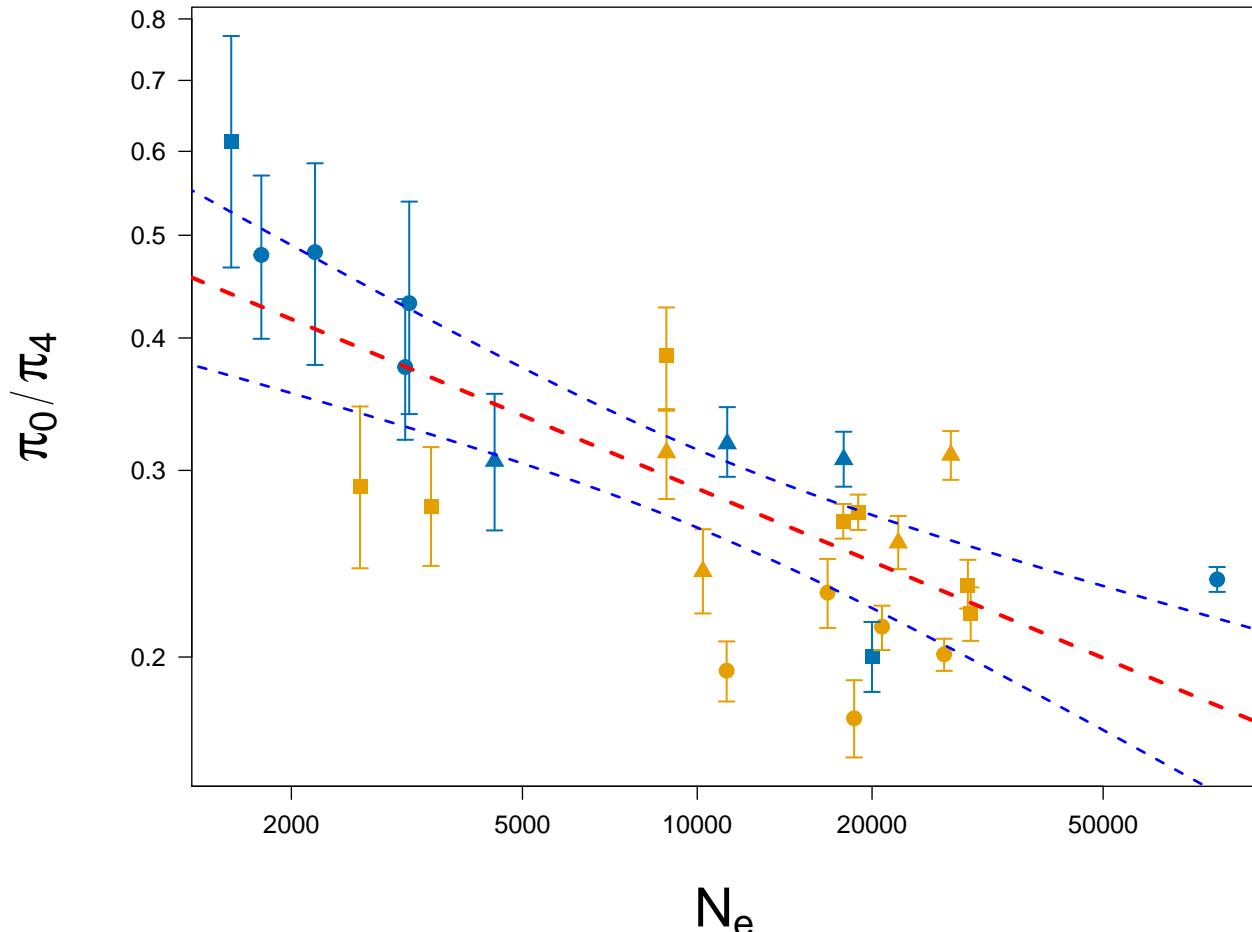
```



```

# summary(model)
par(mar=c(7,7,3,2), mgp=c(5,1,0),cex.lab=2.5,cex.axis=1.2, font=2, las=1)
plot(data_family_daf$Ne_harmonic, data_family_daf$pi0_pi4,log='xy', xlab=expression(N[e]),ylab=expression(pi0_pi4),
pch=ifelse(data_family_daf$sex=='dioecy',19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'),cex=1.5, ylim=c(min(pi0_pi4_lw),max(pi0_pi4_hi)),
# text(data_family_daf$Ne_harmonic, data_family_daf$pi0_pi4, rownames(data_family_daf), pos=3, offset=.5, cex=1.5),
plot_ci(data_family_daf$Ne_harmonic, pi0_pi4_lw, pi0_pi4_hi, col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00')),
# text(data_family_daf$Ne_harmonic, data_family_daf$pi0_pi4, rownames(data_family_daf), pos=3, offset=.5, cex=1.5),
model<-lm(log10(data_family_daf$pi0_pi4)~log10(data_family_daf$Ne_harmonic))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$Ne_harmonic), model, log='xy',lwd=2,lty=2,col='blue')

```



```
m1=pglmm_compare(log10(pi4)~ log10(distSites),phy=tree,data=data_family_daf,REML=F)
m2=pglmm_compare(log10(pi4)~ log10(census),phy=tree,data=data_family_daf,REML=F)
m3=pglmm_compare(log10(pi4)~ log10(Ne_harmonic),phy=tree,data=data_family_daf,REML=F)
R2(m1)
```

```
##      R2_lik   R2_resid   R2_pred
## 0.2070666 0.2070666 0.0000000
```

```
R2(m2)
```

```
##      R2_lik   R2_resid   R2_pred
## 0.4332783 0.4332783 0.0000000
```

```
R2(m3)
```

```
##          R2_lik      R2_resid      R2_pred
## 9.006902e-01 9.006902e-01 -2.220446e-16
```

```
m1=pglmm_compare(log10(pi0_pi4)~ log10(distSites),phy=tree,data=data_family_daf,REML=F)
m2=pglmm_compare(log10(pi0_pi4)~ log10(census),phy=tree,data=data_family_daf,REML=F)
m3=pglmm_compare(log10(pi0_pi4)~ log10(Ne_harmonic),phy=tree,data=data_family_daf,REML=F)
R2(m1)
```

```
##      R2_lik   R2_resid   R2_pred
## 0.24078856 0.29582931 0.03895069
```

```

R2(m2)

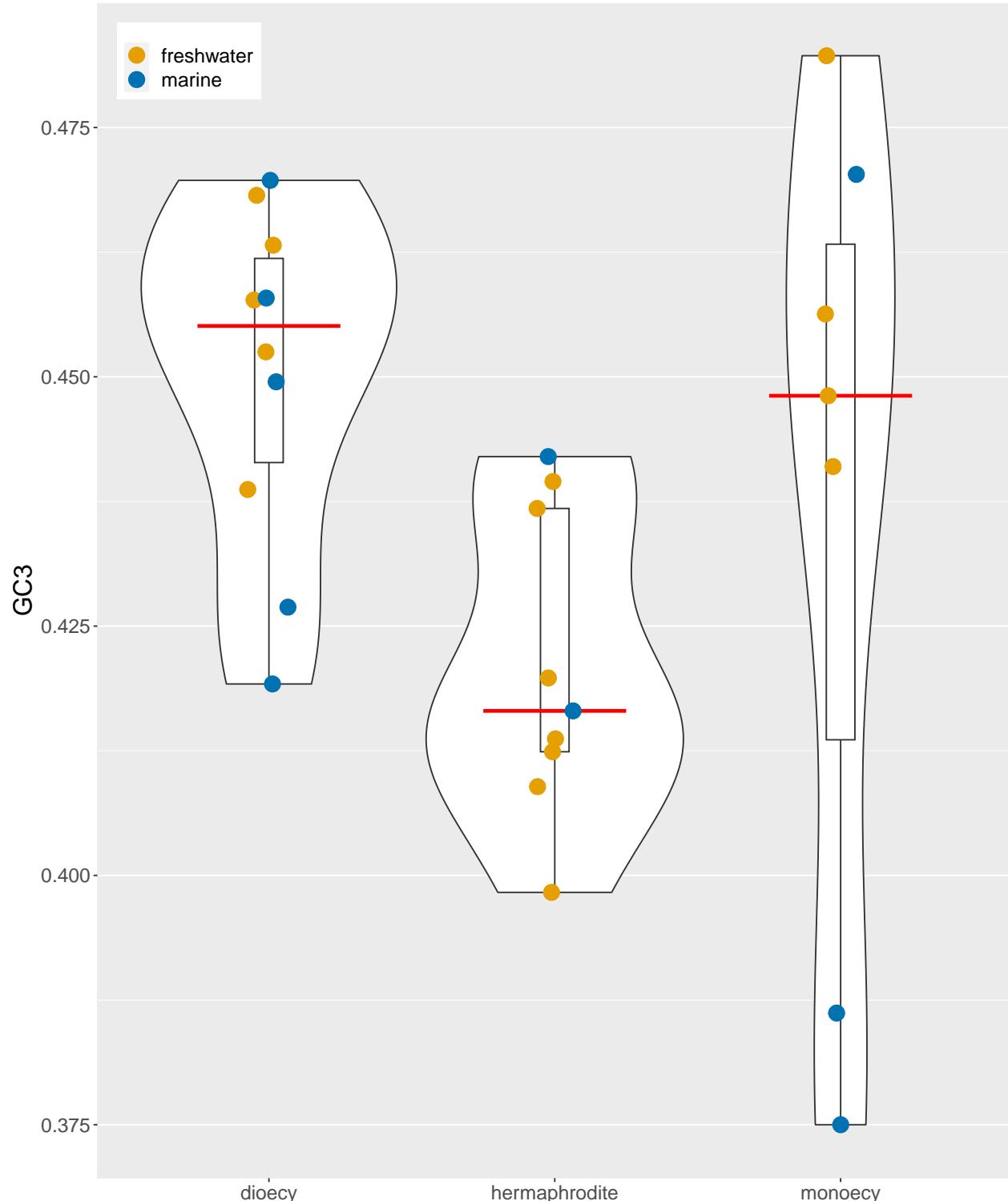
##      R2_lik      R2_resid      R2_pred
## 4.987078e-01 4.987078e-01 -2.220446e-16

R2(m3)

##      R2_lik      R2_resid      R2_pred
## 0.6339442 0.6493877 0.2151063

##### check GC content, genome size and other factors
p<- ggplot(aes(x = sex, y = GC3), data=data_family_daf) +
  geom_violin() +
  geom_boxplot(width=0.1)+ stat_summary(fun = "median", geom = "crossbar",
  width = 0.5, colour = "red") +
  geom_point(aes(fill = habitat, color=habitat), size = 5, shape = 21,
  position = position_jitterdodge(dodge.width=0.1, jitter.width=0.1)) +
  theme(text = element_text(size = 18), axis.title.x = element_blank(),
  panel.grid.minor.x = element_blank(),
  panel.grid.major.x = element_blank())+
  ylab('GC3')
p + scale_fill_manual(values=c('marine'= '#0072B2', 'freshwater' = '#E69F00'))+
  scale_color_manual(values=c('marine'= '#0072B2', 'freshwater' = '#E69F00'))+
  theme(legend.position=c(0.1, 0.95), legend.title = element_blank())

```



```
# hermaphrodite has significant lower GC3 but not after correcting for phylogeny
tapply(data_family_daf$GC3, list(data_family_daf$habitat, data_family_daf$sex), median)
```

```
##          dioecy hermaphrodite monoecy
## freshwater 0.4577      0.41370  0.4522
## marine     0.4495      0.42925  0.3862
```

```

pglmm_compare(GC3~ sex,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:GC3 ~ sex
##
## logLik      AIC      BIC
## 68.92 -127.84 -127.28
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2     1.187e-03 0.034453
## residual 5.702e-05 0.007551
##
## Fixed effects:
##             Value Std.Error Zscore Pvalue
## (Intercept) 0.4599420 0.0076514 60.1125 <2e-16 ***
## sexhermaphrodite -0.0157687 0.0100251 -1.5729 0.1157
## sexmonoecy      -0.0050619 0.0078216 -0.6472 0.5175
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## no different in GC3 between marine and freshwater
tapply(data_family_daf$GC3, list(data_family_daf$habitat),median)

```

```

## freshwater      marine
## 0.44025      0.43445
pglmm_compare(GC3~ habitat,phy=tree,data=data_family_daf,REML=F)

```

```

## Linear mixed model fit by maximum likelihood
##
## Call:GC3 ~ habitat
##
## logLik      AIC      BIC
## 67.76 -127.52 -127.07
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2     1.331e-03 0.036485
## residual 5.823e-05 0.007631
##
## Fixed effects:
##             Value Std.Error Zscore Pvalue
## (Intercept) 0.4581227 0.0078897 58.0656 <2e-16 ***
## habitatmarine -0.0015879 0.0102878 -0.1543 0.8773
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# GC or GC3 has no effects on diversity alone but interact with sex
# pglmm_compare(log10(pi4)~ GC,phy=tree,data=data_family_daf,REML=F)
# pglmm_compare(log10(pi4)~ GC3,phy=tree,data=data_family_daf,REML=F)
pglmm_compare(log10(pi4) ~ GC3*sex,phy=tree,data=data_family_daf,REML=F)

```

```

## Linear mixed model fit by maximum likelihood
##
```

```

## Call:log10(pi4) ~ GC3 * sex
##
## logLik      AIC      BIC
## -11.37   38.73  39.64
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2       0.0000  0.0000
## residual 0.1403  0.3746
##
## Fixed effects:
##             Value Std.Error Zscore    Pvalue
## (Intercept) -16.7143   3.2916 -5.0778 3.818e-07 ***
## GC3          31.0286   7.3044  4.2480 2.157e-05 ***
## sexhermaphrodite 30.1610   4.9332  6.1138 9.726e-10 ***
## sexmonoecy     14.4740   3.6759  3.9376 8.231e-05 ***
## GC3:sexhermaphrodite -69.5003  11.3793 -6.1076 1.011e-09 ***
## GC3:sexmonoecy     -32.1399   8.2015 -3.9188 8.901e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#### G3 improves the explanation of diversity
m.dio = lm(log10(data_family_daf$pi4[data_family_daf$sex=='dioecy'])~data_family_daf$GC3[data_family_daf$sex=='dioecy'])
m.mon = lm(log10(data_family_daf$pi4[data_family_daf$sex=='monoecy'])~data_family_daf$GC3[data_family_daf$sex=='monoecy'])
m.her = lm(log10(data_family_daf$pi4[data_family_daf$sex=='hermaphrodite'])~data_family_daf$GC3[data_family_daf$sex=='hermaphrodite'])

par(mar=c(6,5,3,2), mfrow=c(2,3))
plot(data_family_daf$GC3[data_family_daf$sex=='dioecy'], data_family_daf$pi4[data_family_daf$sex=='dioecy'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='dioecy']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16, xlab='GC3', ylab=expression(pi[4]), cex.lab=2, las=1)
abline(m.dio, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='dioecy']), m.dio, col='blue', log='y', lty=2)
legend('topleft', legend=c('marine', 'freshwater'), pch=16, col=c('#0072B2', '#E69F00'), cex=2)
# text(0.42, 0.02, 'Dioecious', cex=2)

plot(data_family_daf$GC3[data_family_daf$sex=='monoecy'], data_family_daf$pi4[data_family_daf$sex=='monoecy'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='monoecy']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16, xlab='GC3', ylab=expression(pi[4]), cex.lab=2, las=1)
abline(m.mon, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='monoecy']), m.mon, col='blue', log='y', lty=2)
# legend('bottomright', title='Monoecious', legend=c('marine', 'freshwater'), pch=16, col=c('#0072B2',
# text(0.396, 0.0038, 'Monoecious', cex=2)

plot(data_family_daf$GC3[data_family_daf$sex=='hermaphrodite'], data_family_daf$pi4[data_family_daf$sex=='hermaphrodite'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='hermaphrodite']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16, xlab='GC3', ylab=expression(pi[4]), cex.lab=2, las=1)
abline(m.her, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='hermaphrodite']), m.her, col='blue', log='y', lty=2)
# legend('bottomright', title='Hermaphroditic', legend=c('marine', 'freshwater'), pch=16, col=c('#0072B2',
# text(0.396, 0.0038, 'Hermaphroditic', cex=2)

## pi0/pi4

```

```

pglmm_compare(log10(pi0_pi4) ~ GC3*sex,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:log10(pi0_pi4) ~ GC3 * sex
##
## logLik      AIC      BIC
##  20.86 -25.73 -24.82
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.033234 0.18230
## residual 0.004395 0.06629
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) 1.48329  1.08234  1.3705 0.170546
## GC3         -4.64810  2.36916 -1.9619 0.049772 *
## sexhermaphrodite -4.67202  1.42143 -3.2868 0.001013 **
## sexmonoecy    -2.24989  0.99205 -2.2679 0.023334 *
## GC3:sexhermaphrodite 10.73468  3.27280  3.2800 0.001038 **
## GC3:sexmonoecy     5.04559  2.22320  2.2695 0.023237 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

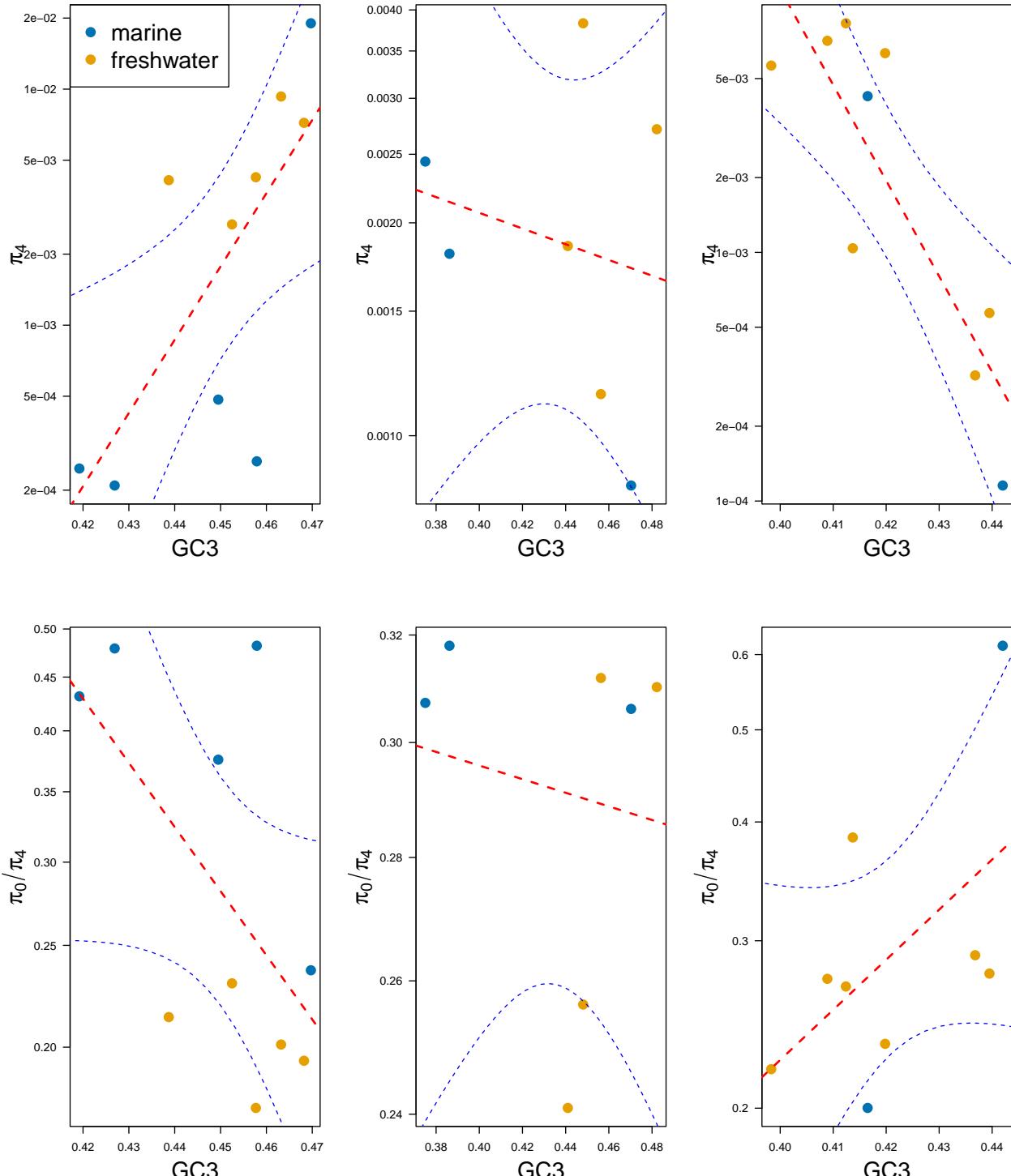
m.dio = lm(log10(data_family_daf$pi0_pi4[data_family_daf$sex=='dioecy'])~data_family_daf$GC3[data_family_daf$sex=='dioecy'])
m.mon = lm(log10(data_family_daf$pi0_pi4[data_family_daf$sex=='monoecy'])~data_family_daf$GC3[data_family_daf$sex=='monoecy'])
m.her = lm(log10(data_family_daf$pi0_pi4[data_family_daf$sex=='hermaphrodite'])~data_family_daf$GC3[data_family_daf$sex=='hermaphrodite'])

plot(data_family_daf$GC3[data_family_daf$sex=='dioecy'], data_family_daf$pi0_pi4[data_family_daf$sex=='dioecy'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='dioecy']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16,xlab='GC3', ylab=expression(pi[0]/pi[4]),cex.lab=2,las=1)
abline(m.dio, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='dioecy']), m.dio, col='blue',log='y',lty=2)

plot(data_family_daf$GC3[data_family_daf$sex=='monoecy'], data_family_daf$pi0_pi4[data_family_daf$sex=='monoecy'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='monoecy']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16,xlab='GC3', ylab=expression(pi[0]/pi[4]),cex.lab=2,las=1)
abline(m.mon, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='monoecy']), m.mon, col='blue',log='y',lty=2)

plot(data_family_daf$GC3[data_family_daf$sex=='hermaphrodite'], data_family_daf$pi0_pi4[data_family_daf$sex=='hermaphrodite'],
      col=ifelse(data_family_daf$habitat[data_family_daf$sex=='hermaphrodite']=='marine', '#0072B2', '#E69F00'),
      log='y', cex=2, pch=16,xlab='GC3', ylab=expression(pi[0]/pi[4]),cex.lab=2,las=1)
abline(m.her, lwd=2, lty=2, col='red')
ci.lines((data_family_daf$GC3[data_family_daf$sex=='hermaphrodite']), m.her, col='blue',log='y',lty=2)

```



```
model=pglmm_compare(log10(pi4) ~ GC3*sex+habitat, phy=tree, data=data_family_daf, REML=F)
R2(model)
```

```
##          R2_lik      R2_resid      R2_pred
##  6.705240e-01  6.705240e-01 -2.220446e-16
model2=pglmm_compare(log10(pi0_pi4) ~ GC3*sex+habitat, phy=tree, data=data_family_daf, REML=F)
R2(model2)
```

```

##          R2_lik      R2_resid      R2_pred
## 5.399969e-01 5.399969e-01 2.220446e-16
## GC or GC3 has a negative effect on Tajima's D

# strong effect of TD on pi0/pi4, especially TDO
pglmm_compare((pi0_pi4)~ TD4,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:(pi0_pi4) ~ TD4
##
## logLik      AIC      BIC
## 24.25 -40.50 -40.05
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.036883 0.19205
## residual 0.001772 0.04209
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) 0.260628 0.043798 5.9506 2.671e-09 ***
## TD4        -0.094268 0.048641 -1.9380 0.05262 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
pglmm_compare((pi0_pi4)~ TDO,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:(pi0_pi4) ~ TDO
##
## logLik      AIC      BIC
## 25.98 -43.97 -43.52
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.020494 0.14316
## residual 0.003279 0.05726
##
## Fixed effects:
##             Value Std.Error Zscore   Pvalue
## (Intercept) 0.272458 0.037322 7.3002 2.873e-13 ***
## TDO        -0.147299 0.049732 -2.9619 0.003058 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## correlated with alpha or strongly positive selected mutations but not with p_b or S_b, neither with str_del
pglmm_compare(str_del ~ GC3*sex+habitat,phy=tree,data=data_family_daf,REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:str_del ~ GC3 * sex + habitat
##
## logLik      AIC      BIC

```

```

## 26.27 -34.53 -33.51
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.000000 0.00000
## residual 0.007764 0.08811
##
## Fixed effects:
##                Value Std.Error Zscore Pvalue
## (Intercept) -0.320241  0.804694 -0.3980 0.690655
## GC3          2.305367  1.773943  1.2996 0.193748
## sexhermaphrodite 3.396252  1.216682  2.7914 0.005248 **
## sexmonoecy    1.004970  0.866757  1.1595 0.246269
## habitatmarine -0.122158  0.040727 -2.9994 0.002705 **
## GC3:sexhermaphrodite -8.065482  2.796871 -2.8838 0.003930 **
## GC3:sexmonoecy     -2.332415  1.932794 -1.2068 0.227525
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#### selfing rate
tapply(data_family_daf$selfingRate, list(data_family_daf$habitat), median)

## freshwater      marine
## 0.17471      0.05308

pglmm_compare(`S_b` ~ selfingRate, phy=tree, data=data_family_daf, REML=F)

## Linear mixed model fit by maximum likelihood
##
## Call:S_b ~ selfingRate
##
## logLik   AIC   BIC
## -108.3  224.5  225.0
##
## Phylogenetic random effects variance (s2):
##          Variance Std.Dev
## s2      0.0      0.0
## residual 242.4   15.57
##
## Fixed effects:
##                Value Std.Error Zscore Pvalue
## (Intercept) 11.6892   4.4285  2.6395 0.0083022 **
## selfingRate 67.1862  18.7314  3.5868 0.0003347 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Ka/Ks
data_family_daf$`Ka/Ks` = as.numeric(data_family_daf$`Ka/Ks`)

## Warning: NAs introduced by coercion
par(mar=c(7,7,3,2), mgp=c(5,1,0), cex.lab=2, cex.axis=1.2, font=4, las=1)

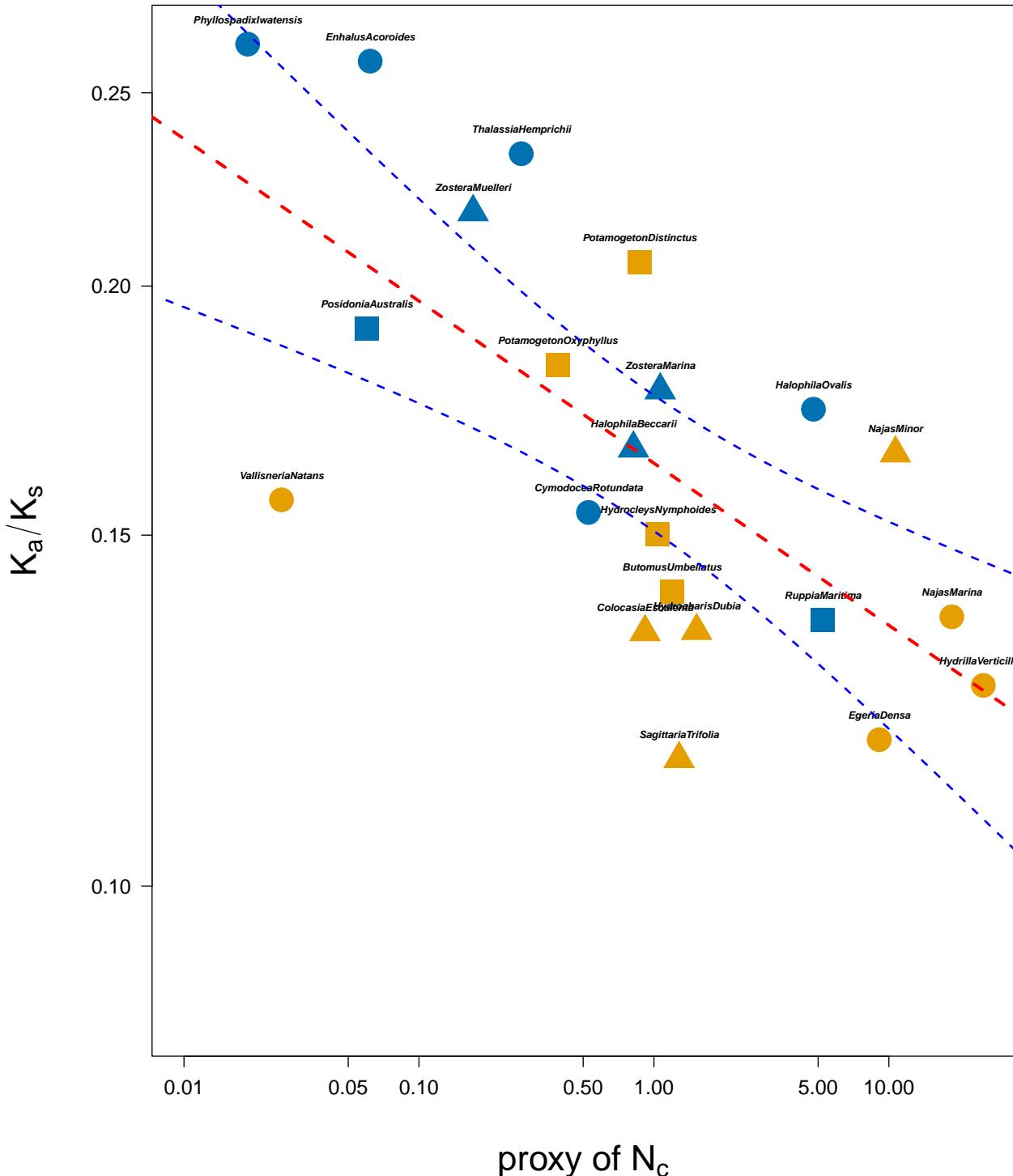
plot(data_family_daf$census, data_family_daf$`Ka/Ks`, log='xy', xlab=expression(paste('proxy of ', N[c])),
      pch=ifelse(data_family_daf$sex=='dioecy', 19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
      col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'), cex=3, xlim=c(0.01,25))
text(data_family_daf$census, data_family_daf$`Ka/Ks`, rownames(data_family_daf), pos=3, offset=1, cex=0.8)

```

```

model<-lm(log10(data_family_daf$`Ka/Ks`)\~log10(data_family_daf$census))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$census), model, log='xy',lwd=2,lty=2,col='blue')

```



```

plot(data_family_daf$Ne_harmonic, data_family_daf$`Ka/Ks`,log='xy', xlab=expression(N[e]),ylab=expression(Ka/Ks),
      pch=ifelse(data_family_daf$sex=='dioecy',19, ifelse(data_family_daf$sex=='monoecy', 17, 15)),
      col=ifelse(data_family_daf$habitat=='marine', '#0072B2', '#E69F00'),cex=3, xlim=c(1e3,1e5))

```

```

text(data_family_daf$Ne_harmonic, data_family_daf$'Ka/Ks', rownames(data_family_daf), pos=3, offset=1, cex=0.8)
model<-lm(log10(data_family_daf$`Ka/Ks`)\~log10(data_family_daf$Ne_harmonic))
abline(model, lwd=3, lty=2, col='red')
ci.lines(log10(data_family_daf$Ne_harmonic), model, log='xy', lwd=2, lty=2, col='blue')

```

