R scripts

1) Ranges of organismic complexity vs. Tmax in groups

```
# complexity ranges vs. Tmax Ranges in groups
Cfrom=c(24,13,4,4)
Cto=c(25,17,8,4)
Tfrom=c(40,42,40,121)
Tto=c(40,54,90,122)
plot(0,col=0,xlim=c(40,122),ylim=c(4,25),xlab="Tmax for growth",ylab="Level of Organismic Complexity")
Tm=c(40 ,40 ,42  ,43  ,52,  54  ,54,  40  ,50,  60,  80,  90  ,121,  122)
C=c(24, 25,  14,  17,  14,  13,  13  ,7  ,8  ,5  ,4  ,6  ,4  ,4)
points(Tm,C,col=2,pch=19)

for (i in 1:4)
lines(c(Tfrom[i],Tto[i]),rep(mean(c(Cfrom[i],Cto[i])),2),lwd=2)
for (i in 1:4)
lines(rep(mean(c(Tfrom[i],Tto[i])),2),c(Cfrom[i],Cto[i]),lwd=2)

x=  apply(rbind(Tfrom,Tto),2,mean)
y=  apply(rbind(Cfrom,Cto),2,mean)
#points(x,y,pch=19)
xx=40:120
points(xx,(xx/84)^(-4)+4,type="l",lty=2,col=3)
# End of complexity ranges vs. Tmax Ranges in Groups
```

2) Tmax analyses for Multicellular Eukarya, Unicellular Eukarya, Bacetria, Archaea

```
# preparative steps for the analyses
# Eukarya
read.table("Multicell_Eukarya.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d1
names(d1)
d1$Domain[grep("Eukarya",d1$Domain)]<-"Muticell Eukarya"
d1$Phylum[grep("Ochro",d1$Phylum)]<-"M-Ochrophyta"
read.table("Unicell_Eukarya.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d2
names(d2)
d2$Domain[grep("Eukarya",d2$Domain)]<-"Unicell Eukarya"
d2$Phylum[grep("Ochro",d2$Phylum)]<-"U-Ochrophyta"

# Bacteria and Archaea
read.table("Bacteria.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d3
names(d3)
read.table("Archaea.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d4
names(d4)
# end of data preparation

# Tmax

# for all oxygen requirements
d<-rbind(d1,d2,d3,d4)
d$Habigroup[d$Habitat %in% c("pelagic","benthic")] <- "nonhydrothermal"
d$Habigroup[d$Habitat %in% c("hydrothermal")] <- "hydrothermal"
table(d$Habigroup)
```

```
dnon<-d[d$Habigroup =="nonhydrothermal",]
dhyd<-d[d$Habigroup =="hydrothermal",]

pairwise.wilcox.test(dnon$Tmax,factor(dnon$Domain),alternative="less",p.adjust.method="BH")
pairwise.wilcox.test(dhyd$Tmax,factor(dhyd$Domain),alternative="less",p.adjust.method="BH")

# just for aerobic
da <- d[d$Oxygen %in% c("aerobic","strictly aerobic","microaerobic"),]
dim(da)
da$Fac <- as.factor(da$Fac)
table(da$Fac)
pairwise.wilcox.test(da$Tmax,da$Fac,alternative="less",p.adjust.method="BH")

# end of Tmax analyses


# Tmax boxplot according to Habitat hydrothermal-nonhydrothermal

# Multicell Eukarya
d1$Habigroup[d1$Habitat %in% c("pelagic","benthic")] <- "nonhydrothermal"
d1$Habigroup[d1$Habitat %in% c("hydrothermal")] <- "hydrothermal"
table(d1$Habigroup)
median(d1$Tmax[d1$Habigroup=="nonhydrothermal"])
median(d1$Tmax[d1$Habigroup=="hydrothermal"])
max(d1$Tmax[d1$Habigroup=="nonhydrothermal"])
max(d1$Tmax[d1$Habigroup=="hydrothermal"])

tapply(d1$Tmax,factor(d1$Habigroup),median)
order(tapply(d1$Tmax,factor(d1$Habigroup),median)) -> o1
tapply(d1$Tmax,factor(d1$Habigroup),median)[o1]
par(mar=c(3,8,2,2))
at.y<-rep(integer(0),2)
at.y[o1] <- 1:2

pdf("boxplot MULTICELL Habigroups.pdf",width=8,height=4)
par(oma=c(2,6,2,2))
boxplot(Tmax~factor(Habigroup),data=d1,horizontal=TRUE,las=1,xlab="Tmax",at=at.y,axes=FALSE,ylim=c(0,13
0),pch=15)
axis(1,at=c(0,20,40,60,80,100,120),las=1)
axis(2,at=1:2,labels=levels(factor(d1$Habigroup))[o1],lwd=0,las=2,tck=0,pos=0)
dev.off()
# end of graphical output for Multicell Eukarya

# Unicell Eukarya
d2$Habigroup[d2$Habitat %in% c("pelagic","benthic")] <- "nonhydrothermal"
d2$Habigroup[d2$Habitat %in% c("hydrothermal")] <- "hydrothermal"
table(d2$Habigroup)
median(d2$Tmax[d2$Habigroup=="nonhydrothermal"])
median(d2$Tmax[d2$Habigroup=="hydrothermal"])
max(d2$Tmax[d2$Habigroup=="nonhydrothermal"])
max(d2$Tmax[d2$Habigroup=="hydrothermal"])
tapply(d2$Tmax,factor(d2$Habigroup),median)
order(tapply(d2$Tmax,factor(d2$Habigroup),median)) -> o2
tapply(d2$Tmax,factor(d2$Habigroup),median)[o2]
par(mar=c(3,8,2,2))
```

```
at.y<-rep(integer(0),2)
at.y[o2] <- 1:2

pdf("boxplot UNIICELL Habigroups.pdf",width=8,height=4)
par(oma=c(2,6,2,2))
boxplot(Tmax~factor(Habigroup),data=d2,horizontal=TRUE,las=1,xlab="Tmax",at=at.y,axes=FALSE,ylim=c(0,13
0),pch=15)
axis(1,at=c(0,20,40,60,80,100,120),las=1)
axis(2,at=1:2,labels=levels(factor(d1$Habigroup))[o2],lwd=0,las=2,tck=0,pos=0)
dev.off()
# end of graphical output for Unicell Eukarya


# Bacteria
d3$Habigroup[d3$Habitat %in% c("pelagic","benthic")] <- "nonhydrothermal"
d3$Habigroup[d3$Habitat %in% c("hydrothermal")] <- "hydrothermal"
table(d3$Habigroup)
median(d3$Tmax[d3$Habigroup=="nonhydrothermal"])
median(d3$Tmax[d3$Habigroup=="hydrothermal"])
max(d3$Tmax[d3$Habigroup=="nonhydrothermal"])
max(d3$Tmax[d3$Habigroup=="hydrothermal"])
tapply(d3$Tmax,factor(d3$Habigroup),median)
order(tapply(d3$Tmax,factor(d3$Habigroup),median)) -> o3
tapply(d3$Tmax,factor(d3$Habigroup),median)[o3]
par(mar=c(3,8,2,2))
at.y<-rep(integer(0),2)
at.y[o3] <- 1:2

pdf("boxplot BAC Habigroups.pdf",width=8,height=4)
par(oma=c(2,6,2,2))
boxplot(Tmax~factor(Habigroup),data=d3,horizontal=TRUE,las=1,xlab="Tmax",at=at.y,axes=FALSE,ylim=c(0,13
0),pch=15)
axis(1,at=c(0,20,40,60,80,100,120),las=1)
axis(2,at=1:2,labels=levels(factor(d1$Habigroup))[o3],lwd=0,las=2,tck=0,pos=0)
dev.off()
# end of graphical output for Bacteria

# Archaea
d4$Habigroup[d4$Habitat %in% c("pelagic","benthic")] <- "nonhydrothermal"
d4$Habigroup[d4$Habitat %in% c("hydrothermal")] <- "hydrothermal"
table(d4$Habigroup)
median(d4$Tmax[d4$Habigroup=="nonhydrothermal"])
median(d4$Tmax[d4$Habigroup=="hydrothermal"])
max(d4$Tmax[d4$Habigroup=="nonhydrothermal"])
max(d4$Tmax[d4$Habigroup=="hydrothermal"])
tapply(d4$Tmax,factor(d4$Habigroup),median)
order(tapply(d4$Tmax,factor(d4$Habigroup),median)) -> o4
tapply(d4$Tmax,factor(d4$Habigroup),median)[o4]
par(mar=c(3,8,2,2))
at.y<-rep(integer(0),2)
at.y[o4] <- 1:2

pdf("boxplot ARCH Habigroups.pdf",width=8,height=4)
par(oma=c(2,6,2,2))
boxplot(Tmax~factor(Habigroup),data=d4,horizontal=TRUE,las=1,xlab="Tmax",at=at.y,axes=FALSE,ylim=c(0,13
0),pch=15)
```

```
axis(1,at=c(0,20,40,60,80,100,120),las=1)
axis(2,at=1:2,labels=levels(factor(d1$Habigroup))[o4],lwd=0,las=2,tck=0,pos=0)
dev.off()
# end of graphical output for Archaea

# end of Tmax boxplots
```

3) Analyses of thermal windows for Multicellular Eukarya, Unicellular Eukarya, Bacetria, Archaea

```
# preparative steps for the analyses
# Eukarya
read.table("Multicell_Eukarya.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d1
names(d1)
d1$Phylum[grep("Ochro",d1$Phylum)]<-"M-Ochrophyta"
read.table("Unicell_Eukarya.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d2
names(d2)
d2$Phylum[grep("Ochro",d2$Phylum)]<-"U-Ochrophyta"

# Bacteria and Archaea
read.table("Bacteria.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d3
names(d3)
read.table("Archaea.txt",dec=",",sep="\t",header=TRUE,comment="",quote="",stringsAsF=FALSE)->d4
names(d4)
# end of data preparation

# Thermal windows , all groups

d<-rbind(d1,d2,d3,d4)
d$Fac <- rep(c("1","2","3","4"),times=c(dim(d1)[1],dim(d2)[1],dim(d3)[1],dim(d4)[1]) )
table(d$Fac)
d$Fac <- as.factor(d$Fac)

da <- d[!( is.na(d$Tmin)),]
da$Fac <- as.factor(da$Fac)
table(da$Fac)
da$Twin=da$Tmax-da$Tmin

# boxplot thermal windows all groups

# all boxplots Twin

median(da$Twin[da$Fac=="1"])
median(da$Twin[da$Fac=="2"])
median(da$Twin[da$Fac=="3"])
median(da$Twin[da$Fac=="4"])
max(da$Twin[da$Fac=="1"])
max(da$Twin[da$Fac=="2"])
max(da$Twin[da$Fac=="3"])
max(da$Twin[da$Fac=="4"])
table(da$Fac)

tapply(da$Twin,factor(da$Fac),median)
order(tapply(da$Twin,factor(da$Fac),median)) -> o6
o6 <- 4:1
tapply(da$Twin,factor(da$Fac),median)[o6]
par(mar=c(3,8,2,2))
```

```
at.y<-rep(integer(0),4)
at.y[o6] <- 1:4

pdf("boxplot ALL Groups.pdf",width=7,height=4.5)
par(oma=c(2,6,2,2))
boxplot(Twin~factor(Fac),data=da,horizontal=TRUE,
      las=1,xlab="Width of thermal window",at=at.y,axes=FALSE,ylim=c(0,60),pch=15)
axis(1,at=c(0,10,20,30,40,50,60),las=1)
nm=c("Multicellular Eukarya","Unicellular Eukarya","Bacteria","Archaea")
axis(2,at=1:4,labels=nm[o6],lwd=0,las=2,tck=0,pos=0)
dev.off()
# end of boxplot All Groups


# for all oxygen requirements
# pairwise Twin Wilcoxon BH corrected on-sided less
da$Twin=da$Tmax-da$Tmin
pairwise.wilcox.test(da$Twin,da$Fac,alternative="less",p.adjust.method="BH")
```

4) Graphical visualization and analyses for critical oxygen concentrations for Eukarya

```
# resampling analyses of pCrit / size data,
# varying over different thresholds seperating "large" from "small"
#
# data preparation:
d <-
read.table("Eucarya.txt",sep="\t",dec=",",stringsAsFactors=FALSE,na.strings="?",header
=TRUE)
d1 <- d[-
c(which(is.na(d$mean.weightg)),which.max(d$cO2)),c("mean.weightg","cO2","Phylum")
]
names(d1)

# not used: scatterplot show distribution of critical oxygen concentrations:
plot(log10(d1$mean.weightg)~d1$cO2,pch=19)
d1$logweight=log10(d1$mean.weightg)
d2<-d1[-which.max(d1$cO2),]
m=median(d2$logweight)
10^m
(s=sd(d2$logweight)/sqrt(70)*6) # 6 standard error of log weight as interval size to
simulate
10^(m-s)
10^(m+s)
sum(d2$mean.weightg<10^(m-s)) # lowest amount simulated in group small : 10
sum(d2$mean.weightg<10^(m+s)) # highest amount simulated in group small : 46
# end of data preparation

# resampling simulations loop
set.seed(48311934)
count.bs.all<- 0
count.bs <- rep(0,5) # column spec. test counter
L=rep(FALSE,100)
j <- 1
a.sample.l <- a.sample.s <- matrix(0,40,7) # not more than 40 needed
for (i in 1:999)
{
```

```
# filter for new cuts only
w.cut=runif(1,min=10^(m-s),max=10^(m+s)) # uniformly resampled size class
boundary
p <- sum(d2$mean.weightg<w.cut) # lowest amount simulated in group small : 29
if (!L[p])
{
d.small=d2[d2$mean.weightg <= w.cut,]
d.large=d2[d2$mean.weightg > w.cut,]
d.s.bs <- d.small[sample(dim(d.small)[1],replace=TRUE),]
d.l.bs <- d.large[sample(dim(d.large)[1],replace=TRUE),]
# build up statistics: how often are all columns of "large" < all corresp. columns in
"small"
hist(d.s.bs$cO2,breaks=c(0,30,60,90,120,150,300),plot=FALSE)$counts -> h.s
hist(d.l.bs$cO2,breaks=c(0,30,60,90,120,150,300),plot=FALSE)$counts -> h.l
diffinv(h.s)-> acc.s
diffinv(h.l)->acc.l
a.sample.s[j,]= acc.s/max(acc.s)
a.sample.l[j,]= acc.l/max(acc.l)
j <- j+1
# barplot(rbind(acc.s/max(acc.s),acc.l/max(acc.l)),beside=TRUE)
a <- (acc.s[2:6]/max(acc.s)>acc.l[2:6]/max(acc.l))
a + count.bs -> count.bs
all(a) + count.bs.all -> count.bs.all
L[p] <- TRUE
}
cat("\r",i)
}
cat("\n")
count.bs
count.bs.all # how many times were all "small" groups greater than corresponding
"large" groups
sum(L) # how many different runs were evaluated
count.bs/sum(L)*100 # percentage sampled "small">"large"
# to compute error bars, analyse standard deviations of sampled comparisons
a.sample.s[(2:j)-1,]
a.sample.l[(2:j)-1,]
apply(a.sample.s[(2:j)-1,],2,sd) -> sd.s
apply(a.sample.l[(2:j)-1,],2,sd) -> sd.l
# accumul. Barplot for median-cut of logweight
pdf("barplot.pdf",width=8,height=6)
w.cut=10^m # median size class boundary
d.small=d2[d2$mean.weightg <= w.cut,]
d.large=d2[d2$mean.weightg > w.cut,]
d.s.bs <- d.small[sample(dim(d.small)[1],replace=TRUE),]
d.l.bs <- d.large[sample(dim(d.large)[1],replace=TRUE),]
# build up statistics: how often are all columns of "large" < all corresp. columns in
"small"
hist(d.s.bs$cO2,breaks=c(0,30,60,90,120,150,300),plot=FALSE)$counts -> h.s
hist(d.l.bs$cO2,breaks=c(0,30,60,90,120,150,300),plot=FALSE)$counts -> h.l
diffinv(h.s)-> acc.s
diffinv(h.l)->acc.l
rbind(acc.s/max(acc.s),acc.l/max(acc.l))[,-1] -> R
colnames(R) <- c("<30","<60","<90","<120","<150","<190")
barplot(R,beside=TRUE,xlab="critical O2 concentration",ylab="accumulated count
fraction")
# error bars from resampling
```

```
for (i in 2:6)
{lines(x=c(i-1,i-1)*3-1.5, y=c(R[1,i-1]-sd.s[i],R[1,i-1]+sd.s[i]),lwd=2,col=2)
lines(x=c(i-1,i-1)*3-0.5, y=c(R[2,i-1]-sd.l[i],R[2,i-1]+sd.l[i]),lwd=2,col=2)
}
title(paste(sum(L)," different weight cuts around median weight"))
dev.off()
# end of graphical output, incl. error bars from resampling
# data preparation for boxplot sorted by medians of pCrit(O2)
order(tapply(d2$cO2,factor(d2$Phylum),median)) -> o
tapply(d2$cO2,factor(d2$Phylum),median)[o]
par(mar=c(3,8,2,2))
at.y<-rep(integer(0),6)
at.y[o] <- 1:6
# end of data preparation

# boxplot of critical O2 conc. for Eucarya
pdf("boxplot.pdf",width=8,height=6.5)
par(oma=c(2,3,2,2))
boxplot(cO2~factor(Phylum),data=d2[-
which.max(d$cO2),],horizontal=TRUE,las=1,xlab="critical O2
concentration",at=at.y,axes=FALSE,ylim=c(0,210))
axis(1,at=c(0:6*30,210),las=1)
axis(2,at=1:6,labels=levels(factor(d$Phylum))[o],las=2,tck=0,pos=0)
axis(2,at=c(0,9),tck=0,labels=c("",""),pos=0)
lines(x=c(60,60),y=c(0,9),lty=2)
dev.off()
# end of boxplot for Eucarya

# end of script
```