# Set the working directory

setwd("C:/Users/pc/Desktop/world")

library(ape)

library(betapart)

library(picante)

library(cluster)

library(dendextend)

library(kmed)

library(recluster)

library(raster)

library(phytools)

library(rgdal)

library(spatialEco)

library(phyloregion)

library(rgeos)

library(jpfxns)

library(dplyr)

library(rangeBuilder)

library(rangemap)

library(speciesgeocodeR)

library(doParallel)

library(parallel)

library(pbapply)

library(colorspace)

load("C:\\Users\\pc\\Desktop\\world\\1\_progetto\_average")

# Load the shapefile

tuscany <- readOGR(choose.files(), "world\_Mollweide.noANTARTID")

crs(tuscany)<-CRS("+proj=moll")

plot(tuscany)

e <- extent(tuscany)

e[3]<-e[3]-200000

# coerce to a SpatialPolygons object

p <- as(e, 'SpatialPolygons')

crs(p)<-CRS("+proj=moll")

mask<- gDifference(p, tuscany)

pt<-pt2

pt$poly\_shp <-subset(pt$poly\_shp , pt$poly\_shp$richness>1)

#remove small islands canarie

pt$poly\_shp<- pt$poly\_shp[!pt$poly\_shp$grids %in% c("v3636", "v3637", "v3638", "v3816", "v4175"), ]

pt$comm\_dat <- pt$comm\_dat[rownames(pt$comm\_dat) %in% pt$poly\_shp$grids, ]

plot\_swatch(pt$poly\_shp, values=pt$poly\_shp$richness, k=20, leg=10, border=NA)

sparse\_comm<- pt$comm\_dat

subphy<-match\_phylo\_comm(tree, sparse\_comm)$phy

submat<-match\_phylo\_comm(tree, sparse\_comm)$com

bb<-beta\_diss(submat, index.family = "sorensen")

pb<-phylobeta(submat, subphy, index.family = "sorensen")

op<-optimal\_phyloregion(pb[[1]], method="average", k=30)

yp<-phyloregion(pb[[1]], shp= pt$poly\_shp, method= "average", k= op$optima$k)

source("1\_funzioni.per.plottare.txt")

dist.p<-as.dist(yp$region.dist)

c1 <- vegan::metaMDS(dist.p, trace = 0)

vp <- data.frame(hex2RGB(hexcols2(c1))@coords)

vp$r <- vp$R \* 255

vp$g <- vp$G \* 255

vp$b <- vp$B \* 255

vp$cluster <- rownames(vp)

pcoa.p<-cmdscale(yp$region.dist)

rgbcol.p<-recluster.col(pcoa.p)

clusters.p<- (hclust(yp$region.dist, "average"))$order

new\_colours\_sor<-recluster.group.col(rgbcol.p, clusters.p)

newcol.p<- new\_colours\_sor$aggr

newcol.p2<-as.data.frame(new\_colours\_sor$all)

newcol.p2$cluster<-clusters.p

newcol.p2$plot<-rownames(newcol.p2)

nomi<- as.data.frame(clusters.p)

nomi$order<-seq(1, length(clusters.p))

names(nomi)<-c("plot2", "order")

nomi<-nomi[order(nomi$plot2),]

newcol.p2$plot2<-nomi$plot2

tree.p <- read.tree(text = write.tree(ladderize(as.phylo(hclust(yp$region.dist, "average")))))

nomi2<-as.data.frame(as.numeric(tree.p$tip.label))

nomi2$order<-seq(1, length(tree.p$tip.label))

names(nomi2)<-c("plot3", "order2")

newcol.p2<-merge(newcol.p2, nomi2, by.x="plot2", by.y="plot3")

newcol.p2 <-newcol.p2[order(newcol.p2$order2),]

newcol.p2<-unique(newcol.p2[,c(2:6)])

dend.p<-ladderize(as.dendrogram(hclust(yp$region.dist, "average")))

col.p<-rgb(newcol.p2[, 3], newcol.p2[,4], newcol.p2[, 5],maxColorValue=255)

newcol.p2<-vp

newcol.p2<-merge(newcol.p2, nomi2, by.x="cluster", by.y="plot3")

newcol.p2 <-newcol.p2[order(newcol.p2$order2),]

dend.p<-ladderize(as.dendrogram(hclust(yp$region.dist, "average")))

col.p<-rgb(newcol.p2[, 5], newcol.p2[,6], newcol.p2[, 7],maxColorValue=255)

#pdf("res.pd.FINAL.pdf", width=13, height=11, useDingbats = F)

#tiff("res.pd.new.0.25.max30.200.no.hawaii.tiff", units="in", width=13, height=11, res=300)

#setEPS()

#postscript("res.pd.new.eps", width=13, height=11)

layout(mat = matrix(c(1, 2, 1, 3, 1, 4),

nrow = 2,

ncol = 3),

heights = c(2, 1.25), # Heights of the two rows

widths = c(1, 1)) # Widths of the two columns

par(mar = c(0, 1, 5, 7))

par(oma = c(0,0,0,0))

plot(yp, palette="NMDS", cex=2, border="grey")

#plot(yp2, palette="NMDS", cex=2, border=c("black", NA, NA), col=rep(NA, 3), add=T)

plot(mask, add=T, col="white", border=NA)

plot(tuscany, add=T, border="grey")

plot(yp, palette="NMDS", cex=2, border=NA, col=rep(NA, op$optima$k), add=T)

title("(a)", adj = 0.09, line = -6, cex.main=2)

par(mar = c(5, 10, 2, 1))

plot\_NMDS(yp, cex=4, cex.axis=1.5, cex.lab=1.5)

text\_NMDS(yp, cex=1.5)

text(x = -0.2, y = 0.18, cex=1.5, labels = paste("Stress = ",round(yp$NMDS$stress,3)))

title("(b)", adj = 0, line = 0.5, cex.main=2)

par(mar = c(5, 6, 2, 3))

plot(op$df$k, op$df$ev, ylab = "Explained variances", xlab = "Number of clusters", cex=1.5, cex.lab=1.5, cex.axis=1.5, pch=16)

lines(op$df$k[order(op$df$k)], op$df$ev[order(op$df$k)],pch=1)

points(op$optimal$k, op$optimal$ev, pch=21, bg="red", cex=3)

points(op$optimal$k, op$optimal$ev, pch=21, bg="red", type = "h", lty=2)

title("(c)", adj = 0, line = 0.5, cex.main=2)

par(mar = c(5, 2, 2, 10))

dend.p %>% set("labels\_cex", 1.5) %>% set("branches\_k\_color", value= col.p, k= length(clusters.p)) %>% set("branches\_lwd", 7) %>% plot (horiz = F, axes=T, cex.axis=1.5)

title("(d)", adj = 0, line = 0.5, cex.main=2)

#dev.off()

mantel(pb[[1]], bb[[1]])

ssb<-select\_linkage(bb[[1]])

ssp<-select\_linkage(pb[[1]])

par(mfrow=c(1,2))

barplot(ssb, horiz=TRUE, las=1)

barplot(ssp, horiz=TRUE, las=1)

tableS1<-cbind(as.data.frame(ssp), as.data.frame(ssb))

names(tableS1)<-c("pβsim", "βsim")

write.csv(tableS1, "tableS1.csv")

##table2

longF<- sparse2long(pt$comm\_dat)

memb<-yp$membership

comm<-merge(longF, memb, by="grids")

comm<-unique(comm[,c(2,3)])

comm<-long2sparse(comm, grids = "cluster", species = "species")

pd <- PD(comm, subphy)

regions<-list()

for (z in 1:length(yp$shp@polygons)){

list1<-list()

for (i in 1:length(yp$shp@polygons[[z]]@Polygons)){

list1[[i]] <- SpatialPolygons(list(Polygons(list(yp$shp@polygons[[z]]@Polygons[[i]]),1)))

crs(list1[[i]]) <- CRS("+proj=moll")

}

regions[[z]]<-do.call(rbind,c(makeUniqueIDs = TRUE, list1))

}

names(regions) <- names(pd)

areas<-c()

for (i in 1:length(regions)){

areas[[i]] <- sum(area(regions[[i]]))/1000000

}

names(areas) <- names(pd)

ed<-unique(yp$region.df[,c(2,3)])

table2<-data.frame(cbind(areas, pd))

table2<-data.frame(units = rownames(table2), table2)

comm<- sparse2dense(comm)

sr<- rowSums(comm)

names(sr)<-rownames(comm)

table2$sr<-sr

table2$units<-as.numeric(as.character(table2$units))

table2 <- table2[order(table2$units),]

table2$ed<-ed[,2]

write.csv(table2, "table2b.csv", row.names=FALSE)

###table2 end

###shape file

kindgoms.shape <- merge(pt$poly\_shp, data.frame(grids= yp$region.df$grids, ED= yp$region.df$ED, cluster=yp$region.df$cluster), by="grids")

shp <- aggregate(kindgoms.shape, by = 'cluster')

clip<-crop(shp, tuscany)

writeOGR(clip, ".", "kindgoms.shape", driver="ESRI Shapefile")

###shape file end

##### ED mapping

yp<-phyloregion(pb[[1]], shp= pt$poly\_shp, method= "average", k= 16)

y <- merge(pt$poly\_shp, data.frame(grids= yp$region.df$grids, ED= yp$region.df$ED), by="grids")

y <- y[!is.na(y@data$ED),]

plot\_swatch(y, values = y$ED, k = 30, border=NA, col = hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), main="phylogenetic distinctiveness")

r<-raster(resolution=200000, xmn= -17702609, xmx=17697391, ymn= -6649878, ymx=8750122, crs="+proj=moll")

yr1<-map <- rasterize(y, r, field=y$ED)

plot(yr1, zlim=c(0.45,0.65), main="", col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE))

#####

tree2<-extract.clade(tree, "Magnoliophyta")

subphy2<-match\_phylo\_comm(tree2, sparse\_comm)$phy

submat2<-match\_phylo\_comm(tree2, sparse\_comm)$com

pb2<-phylobeta(submat2, subphy2, index.family = "sorensen")

###ED

x<- pb[[1]]

k=16

method="average"

Q <- as.dist(x)

P1 <- hclust(Q, method = method)

g <- cutree(P1, k)

dx <- data.frame(grids=names(g), cluster = unname(g))

x <- as.matrix(x)

colnames(x) <- rownames(x)

region.mat <- matrix(NA, k, k, dimnames = list(1:k, 1:k))

for (i in 1:k) {

for (j in 1:k) {

region.mat[i, j] <- mean(x[names(g)[g == i], names(g)[g == j]])

}

}

region.dist <- as.dist(region.mat)

region.mat <- as.matrix(region.dist)

evol\_distinct <- colSums(region.mat) / (nrow(region.mat) - 1)

evol\_distinct <- data.frame(ED = evol\_distinct)

evol\_distinct <- cbind(cluster = rownames(evol\_distinct),

data.frame(evol\_distinct, row.names = NULL))

evol\_distinct1<- evol\_distinct

###

x<- pb[[1]]

k=16

method="average"

Q <- as.dist(x)

P1 <- hclust(Q, method = method)

g <- cutree(P1, k)

dx <- data.frame(grids=names(g), cluster = unname(g))

x <- as.matrix(pb2[[1]])

colnames(x) <- rownames(x)

region.mat <- matrix(NA, k, k, dimnames = list(1:k, 1:k))

for (i in 1:k) {

for (j in 1:k) {

region.mat[i, j] <- mean(x[names(g)[g == i], names(g)[g == j]])

}

}

region.dist <- as.dist(region.mat)

region.mat <- as.matrix(region.dist)

evol\_distinct <- colSums(region.mat) / (nrow(region.mat) - 1)

evol\_distinct <- data.frame(ED = evol\_distinct)

evol\_distinct <- cbind(cluster = rownames(evol\_distinct),

data.frame(evol\_distinct, row.names = NULL))

evol\_distinct2<- evol\_distinct

e1<-merge(yp$region.df, evol\_distinct1, by="cluster")

y <- merge(pt$poly\_shp, data.frame(grids= e1$grids, ED= e1$ED.y), by="grids")

y <- y[!is.na(y@data$ED),]

plot\_swatch(y, values = y$ED, k = 30, border=NA, col = hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), main="phylogenetic distinctiveness")

r<-raster(resolution=200000, xmn= -17702609, xmx=17697391, ymn= -6649878, ymx=8750122, crs="+proj=moll")

yr1<-map <- rasterize(y, r, field=y$ED)

plot(yr1, zlim=c(0.45,0.65), main="", col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE))

e2<-merge(yp$region.df, evol\_distinct2, by="cluster")

y2 <- merge(pt$poly\_shp, data.frame(grids= e2$grids, ED= e2$ED.y), by="grids")

y2 <- y2[!is.na(y2@data$ED),]

plot\_swatch(y2, values = y2$ED, k = 30, border=NA, col = hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), main="phylogenetic distinctiveness")

yr2<-map <- rasterize(y2, r, field=y2$ED)

plot(yr2, zlim=c(0.45,0.65), main="", col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE))

tiff("res.ED1.tiff", units="in", width=13, height=10, res=300)

par(mar = c(0, 0, 0, 0))

plot(yr1, zlim=c(0.45,0.65), col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), axes=F, box=F, legend=F)

plot(mask, add=T, col="white", border=NA)

plot(tuscany, add=T, border="grey")

title("(a)", adj = 0.08, line = -12, cex.main=2.5)

dev.off()

tiff("res.ED2.tiff", units="in", width=13, height=10, res=300)

par(mar = c(0, 0, 0, 0))

plot(yr2, zlim=c(0.45,0.65), col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), axes=F, box=F, legend=F)

plot(mask, add=T, col="white", border=NA)

plot(tuscany, add=T, border="grey")

title("(b)", adj = 0.08, line = -12, cex.main=2.5)

dev.off()

tiff("res.ED.legend.tiff", units="in", width=13, height=11, res=300)

par(mar = c(5, 5, 25, 1))

plot(yr1, zlim=c(0.45,0.65), col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), axes=F, box=F, legend=F)

plot(yr1, legend.only=T, horizontal=T, col=hcl.colors(n=30, palette = "YlOrRd", rev = TRUE), zlim=c(0.45,0.65), legend.width=2)

dev.off()

###ED mapping stop

##### mapping kingdoms for K 3,4,5,6

k<-6

yp<-phyloregion(pb[[1]], shp= pt$poly\_shp, method= "average", k= k)

source("1\_funzioni.per.plottare.txt")

dist.p<-as.dist(yp$region.dist)

c1 <- vegan::metaMDS(dist.p, trace = 0)

vp <- data.frame(hex2RGB(hexcols2(c1))@coords)

vp$r <- vp$R \* 255

vp$g <- vp$G \* 255

vp$b <- vp$B \* 255

vp$cluster <- rownames(vp)

pcoa.p<-cmdscale(yp$region.dist)

rgbcol.p<-recluster.col(pcoa.p)

clusters.p<- (hclust(yp$region.dist, "average"))$order

new\_colours\_sor<-recluster.group.col(rgbcol.p, clusters.p)

newcol.p<- new\_colours\_sor$aggr

newcol.p2<-as.data.frame(new\_colours\_sor$all)

newcol.p2$cluster<-clusters.p

newcol.p2$plot<-rownames(newcol.p2)

nomi<- as.data.frame(clusters.p)

nomi$order<-seq(1, length(clusters.p))

names(nomi)<-c("plot2", "order")

nomi<-nomi[order(nomi$plot2),]

newcol.p2$plot2<-nomi$plot2

tree.p <- read.tree(text = write.tree(ladderize(as.phylo(hclust(yp$region.dist, "average")))))

nomi2<-as.data.frame(as.numeric(tree.p$tip.label))

nomi2$order<-seq(1, length(tree.p$tip.label))

names(nomi2)<-c("plot3", "order2")

newcol.p2<-merge(newcol.p2, nomi2, by.x="plot2", by.y="plot3")

newcol.p2 <-newcol.p2[order(newcol.p2$order2),]

newcol.p2<-unique(newcol.p2[,c(2:6)])

dend.p<-ladderize(as.dendrogram(hclust(yp$region.dist, "average")))

col.p<-rgb(newcol.p2[, 3], newcol.p2[,4], newcol.p2[, 5],maxColorValue=255)

newcol.p2<-vp

newcol.p2<-merge(newcol.p2, nomi2, by.x="cluster", by.y="plot3")

newcol.p2 <-newcol.p2[order(newcol.p2$order2),]

dend.p<-ladderize(as.dendrogram(hclust(yp$region.dist, "average")))

col.p<-rgb(newcol.p2[, 5], newcol.p2[,6], newcol.p2[, 7],maxColorValue=255)

tiff("res.pd.new.k6.tiff", units="in", width=13, height=11, res=300)

par(mfrow=c(2,1))

par(mar = c(1, 5, 5, 5))

plot(yp, palette="NMDS", cex=2, border="grey", main="phylogenetic regionalisation for k = 6")

plot(mask, add=T, col="white", border=NA)

plot(tuscany, add=T, border="grey")

plot(yp, palette="NMDS", cex=2, border=NA, col=rep(NA, k), add=T)

title("(a)", adj = 0.11, line = -3, cex.main=1.5)

par(mar = c(10, 20, 2, 15))

dend.p %>% set("labels\_cex", 1.5) %>% set("branches\_k\_color", value= col.p, k= length(clusters.p)) %>% set("branches\_lwd", 7) %>% plot (horiz = F, axes=T, cex.axis=1.5)

title("(b)", adj = 0, line = 0.5, cex.main=1.5)

dev.off()

##### mapping taxonomic dissimilarity

ob<-optimal\_phyloregion(bb[[1]], method="average", k=30)

yb<-phyloregion(bb[[1]], shp= pt$poly\_shp, method= "average", k= ob$optima$k)

tax.shp <- merge(pt$poly\_shp, data.frame(grids= yb$region.df$grids, ED= yb$region.df$ED, cluster=yb$region.df$cluster), by="grids")

ptt<-pt

ptt$poly\_shp<- ptt$poly\_shp[!ptt$poly\_shp$grids %in% c("v652", "v653"), ]

ptt$comm\_dat <- ptt$comm\_dat[rownames(ptt$comm\_dat) %in% ptt$poly\_shp$grids, ]

plot\_swatch(ptt$poly\_shp, values=ptt$poly\_shp$richness, k=20, leg=10, border=NA)

sparse\_commT<- ptt$comm\_dat

subphyT<-match\_phylo\_comm(tree, sparse\_commT)$phy

submatT<-match\_phylo\_comm(tree, sparse\_commT)$com

bb<-beta\_diss(submatT, index.family = "sorensen")

ob<-optimal\_phyloregion(bb[[1]], method="average", k=30)

yb<-phyloregion(bb[[1]], shp= ptt$poly\_shp, method= "average", k= ob$optima$k)

#########

dist.t<-as.dist(yb$region.dist)

c1 <- vegan::metaMDS(dist.t, trace = 0)

vt <- data.frame(hex2RGB(hexcols2(c1))@coords)

vt$r <- vt$R \* 255

vt$g <- vt$G \* 255

vt$b <- vt$B \* 255

vt$cluster <- rownames(vt)

pcoa.t<-cmdscale(yb$region.dist)

rgbcol.t<-recluster.col(pcoa.t)

clusters.t<- (hclust(yb$region.dist, "average"))$order

new\_colours\_sor<-recluster.group.col(rgbcol.t, clusters.t)

newcol.t<- new\_colours\_sor$aggr

newcol.t2<-as.data.frame(new\_colours\_sor$all)

newcol.t2$cluster<-clusters.t

newcol.t2$plot<-rownames(newcol.t2)

nomi<- as.data.frame(clusters.t)

nomi$order<-seq(1, length(clusters.t))

names(nomi)<-c("plot2", "order")

nomi<-nomi[order(nomi$plot2),]

newcol.t2$plot2<-nomi$plot2

tree.t <- read.tree(text = write.tree(ladderize(as.phylo(hclust(yb$region.dist, "average")))))

nomi2<-as.data.frame(as.numeric(tree.t$tip.label))

nomi2$order<-seq(1, length(tree.t$tip.label))

names(nomi2)<-c("plot3", "order2")

newcol.t2<-merge(newcol.t2, nomi2, by.x="plot2", by.y="plot3")

newcol.t2 <-newcol.t2[order(newcol.t2$order2),]

newcol.t2<-unique(newcol.t2[,c(2:6)])

dend.t<-ladderize(as.dendrogram(hclust(yb$region.dist, "average")))

col.t<-rgb(newcol.t2[, 3], newcol.t2[,4], newcol.t2[, 5],maxColorValue=255)

newcol.t2<-vt

newcol.t2<-merge(newcol.t2, nomi2, by.x="cluster", by.y="plot3")

newcol.t2 <-newcol.t2[order(newcol.t2$order2),]

dend.t<-ladderize(as.dendrogram(hclust(yb$region.dist, "average")))

col.t<-rgb(newcol.t2[, 5], newcol.t2[,6], newcol.t2[, 7],maxColorValue=255)

#pdf("res.tx.new0.25.max30.200.pdf", width=17, height=14)

tiff("res.tx.final.tiff", units="in", width=13, height=11, res=300)

# Set plot layout

layout(mat = matrix(c(1, 2, 1, 3, 1, 4),

nrow = 2,

ncol = 3),

heights = c(2, 1.25), # Heights of the two rows

widths = c(1, 1)) # Widths of the two columns

par(mar = c(1, 5, 5, 5))

plot(yb, palette="NMDS", cex=2, border="grey")

plot(mask, add=T, col="white", border=NA)

plot(tuscany, add=T, border="grey")

plot(yb, palette="NMDS", cex=2, border=NA, col=rep(NA, ob$optima$k), add=T)

title("(a)", adj = 0.09, line = -6, cex.main=2)

par(mar = c(5, 10, 2, 1))

plot\_NMDS(yb, cex=4, cex.axis=1.5, cex.lab=1.5)

text\_NMDS(yb, cex=1.5)

text(x = -0.35, y = 0.18, labels = paste("Stress = ",round(yb$NMDS$stress,3)))

title("(b)", adj = 0, line = 0.5, cex.main=2)

par(mar = c(5, 6, 2, 5))

plot(ob$df$k, ob$df$ev, ylab = "Explained variances", xlab = "Number of clusters", cex=1.5, cex.lab=1.5, cex.axis=1.5, pch=16)

lines(ob$df$k[order(ob$df$k)], ob$df$ev[order(ob$df$k)],pch=1)

points(ob$optimal$k, ob$optimal$ev, pch=21, bg="red", cex=3)

points(ob$optimal$k, ob$optimal$ev, pch=21, bg="red", type = "h", lty=2)

title("(c)", adj = 0, line = 0.5, cex.main=2)

par(mar = c(5, 0, 2, 9))

dend.t %>% set("labels\_cex", 1.2) %>% set("branches\_k\_color", value= col.t, k= length(clusters.t)) %>% set("branches\_lwd", 7) %>% plot (horiz = F, axes=T, cex.axis=1.5)

title("(d)", adj = 0, line = 0.5, cex.main=2)

dev.off()