

Curso Intensivo Intersemestral
(Paleo)Bio Indicadores Neotropicales

Métodos estadísticos aplicados a la reconstrucción ambiental

Dr. Alexander Correa
Instituto de Geología, UNAM



MEXIDRILL
CHALCO



POSGRADO EN CIENCIAS DEL MAR Y LIMNOLOGÍA

R: ambiente de trabajo y programación

Principio de trabajo: Todo es un objeto y todo objeto tiene clase (class).

Sintaxis básica: `function(argument1,argument2 ...)`

Indexación: Selección de elementos dentro de los objetos.

Clase	Descripción	Indexación
Matrix	Arreglo bidimensional de datos numéricos (filas, columnas)	A[filas,columna]
Data frame	Arreglo bidimensional de datos (filas, columnas)	A[filas,columna]
Vector character	Arreglo unidimensional de datos	A[posición]
List	Colección de objetos	A\$objeto A[[objeto]]

Funciones básicas

```
load("pi6.pollenp.RData")
```

```
load("pi6.metadata.RData")
```

```
class(pi6.pollenp)
```

```
colnames(pi6.pollenp)
```

```
dim(pi6.pollenp)
```

```
summary(pi6.pollenp)
```

```
mean(pi6.pollenp[,3])
```

```
sd(pi6.pollenp[,3])
```

```
scale(pi6.pollenp)
```

```
which(pi6.pollenp[,103]>40)
```

```
which(pi6.pollenp[,103]>=40)
```

```
which(pi6.pollenp[,103]<5)
```

```
which(pi6.pollenp[,103]=<10)
```

```
which(pi6.pollenp[,104]==0)
```

```
which(pi6.pollenp[,103]!=0)
```

Cuantificación de cambios ambientales y ecológicos: Interpretación *a priori*

- Reconstrucción ambiental y ecológica basada en datos paleo: Información ecológica basada en la interpretación de secuencias fósiles.
- Conocimiento previo sobre ecología de las especies es clave en la interpretación *a priori* de los resultados.
- Herramientas de ordenación orientadas hacia dos objetivos fundamentales: i) Estandarizar interpretaciones, y ii) maximizar la información obtenida a partir del carácter multivariado de las bases de datos.

```
load("pi6.pollenp.RData")
```

```
load("pi6.metadata.RData")
```

```
read.csv("filename.csv")
```

```
Mac: read.table(pipe("pbpaste"),header)
```

```
PC: read.table("clipboard",header)
```

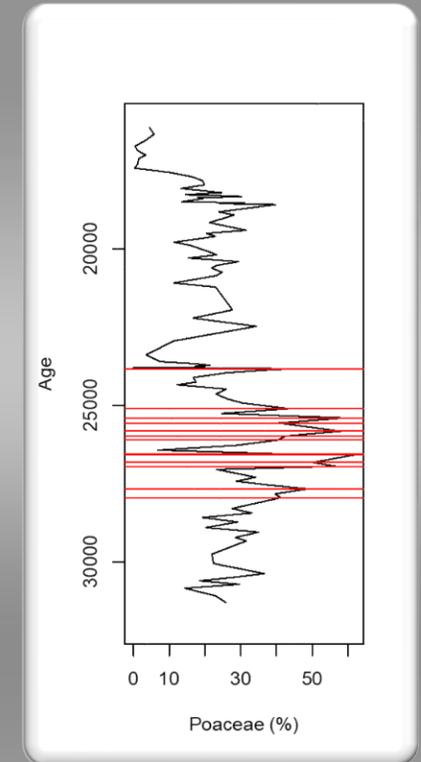
```
apply(pi6.metadata,2,range)
```

```
plot(pi6.pollenp[,103],pi6.metadata[,2],t="l",ylab="Age",  
      xlab="Poaceae (%)",ylim=c(32000,16000))
```

```
which(pi6.pollenp[,103]>40)->a
```

```
pi6.metadata[a,2]
```

```
abline(h=pi6.metadata[a,2],col="red")
```



Métodos de ordenación multivariados

- Bases de datos se componen de sitios (muestras) y sus descriptores (variables = especies), organizados matricialmente en filas y columnas, respectivamente.
- Métodos multivariados resumen la variabilidad a través de la integración de múltiples descriptores (variables = especies) a través de ejes que resumen la variabilidad de los datos (PCA, CA, PCoA, MDS, NMDS, etc.).
- Paquete vegan de R ofrece la implementación de muchos de estos métodos.

```
colnames(pi6.pollenp)
```

```
dim(pi6.pollenp)
```

```
apply(pi6.pollenp,2,range)
```

```
install.packages("vegan",dependencies)
```

```
library(vegan)
```

- Análisis de Correspondencia sin tendencia (DCA) ha probado ser de gran utilidad para el análisis de bases de datos de comunidades fósiles.
- Se asume que la estructura y composición de las comunidades fósiles están controladas por factores ambientales: la variabilidad ambiental expresada en los ejes identificados por la ordenación está asociada a la variabilidad del ambiente.

?decorana

- DCA: Estandarización sucesiva de filas y columnas, hasta encontrar sumas convergentes (en filas y columnas).
`decorana(pi6.pollenp[,3:146])->dca.pi`
`summary(dca.pi)`
- Las sumas convergentes de filas y columnas representan los puntajes de la ordenación para muestras y especies, respectivamente.
`scores(dca.pi,display="species")->scores.spp`
`install.packages("paleoMAS",dependencies)`
`library(paleoMAS)`
- Soluciona el efecto de arco producido por otras técnicas mediante la segmentación de la base de datos.
`colnames(filter.p(pi6.pollenp[,3:146])$filtered)->spp`
- La segmentación corrompe la varianza de manera que valores propios (eigenvalues) NO representan varianza.
`scores.spp[spp,1:2]->scores.spp1`
`plot(scores.spp1[,1:2],asp=1,pch=16,cex=0.7)`
`abline(h=0,v=0,lty=2,col="grey")`
`text(scores.spp1[,1:2],labels=spp,cex=0.8,pos=3)`

- La interpretación de los resultados se da a posteriori a través de la identificación de los gradientes representados por los ejes de la ordenación mediante conocimiento de experto.
- Se pueden proyectar vectores de indicadores asociados con variables ambientales independientes que ayuden en la identificación de los gradientes ambientales asociados a los ejes.

```
par(mfrow=c(1,2))
```

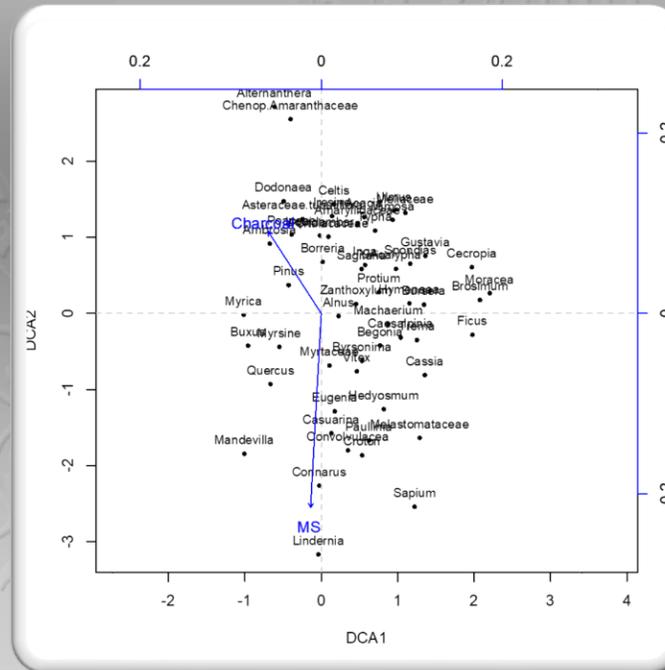
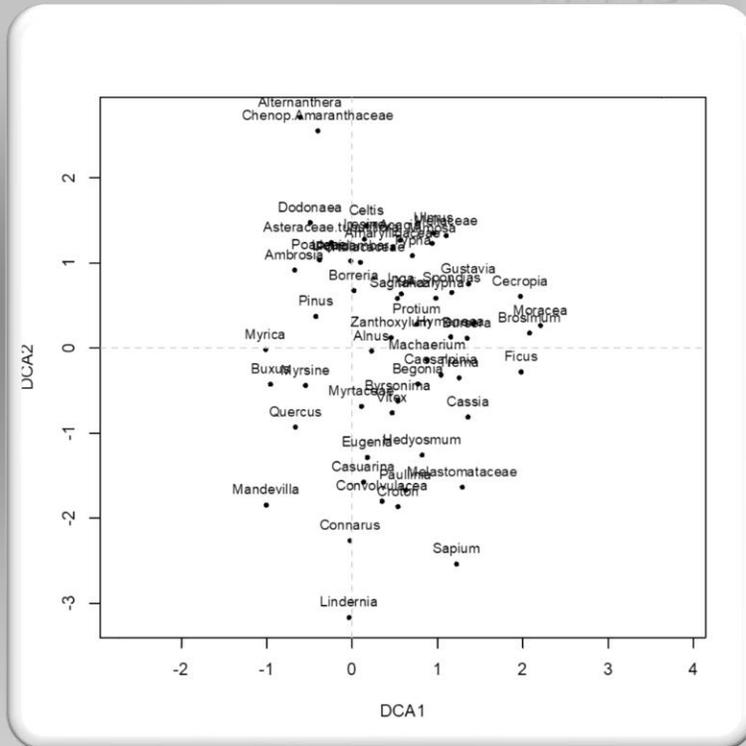
```
plot(scores.spp1[,1:2],asp=1,pch=16,cex=0.7)
```

```
abline(h=0,v=0,lty=2,col="grey")
```

```
text(scores.spp1[,1:2],labels=spp,cex=0.8,pos=3)
```

```
envfit(dca.pi,pi6.metadata[,3:4])->independent
```

```
plot(independent,axis=T,labels=c("MS","Charcoal"))
```



- Las especies definen el espacio ecológico y/o ambiental en el que se ubican las muestras.
- Así, la longitud de los ejes de las muestras siempre va a ser menor que la longitud de los ejes de las especies.

```
par(mfrow=c(1,2))
```

```
plot(scores.spp1[,1:2],asp=1,pch=16,cex=0.7)
```

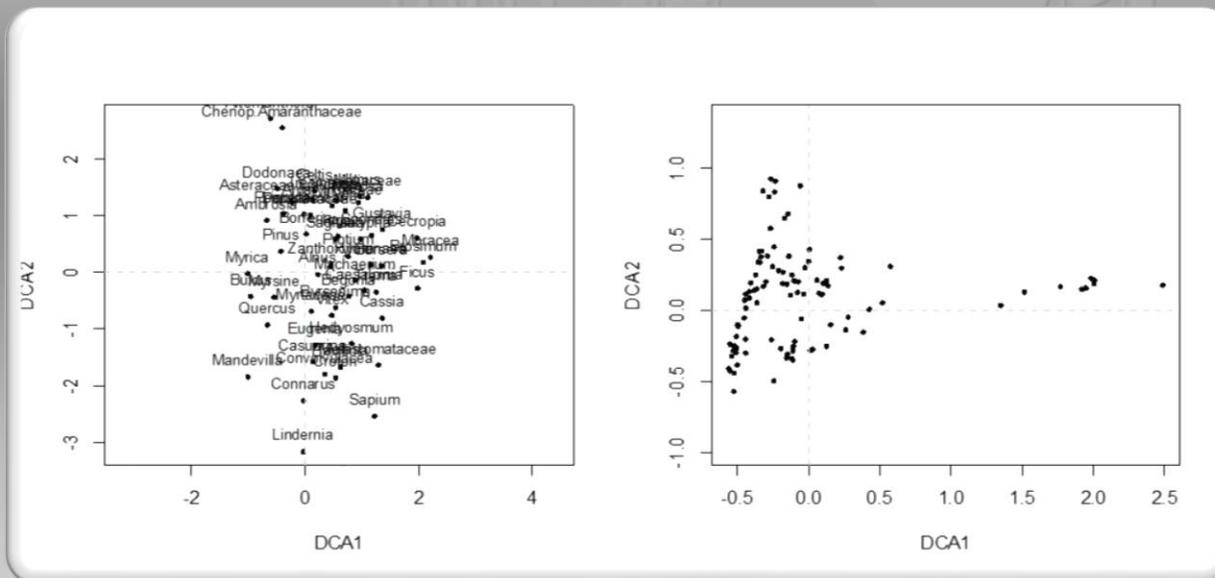
```
abline(h=0,v=0,lty=2,col="grey")
```

```
text(scores.spp1[,1:2],labels=spp,cex=0.8,pos=3)
```

```
scores(dca.pi,display="sites")->scores.samples
```

```
plot(scores.samples[,1:2],asp=1,pch=16,cex=0.7)
```

```
abline(h=0,v=0,lty=2,col="grey")
```

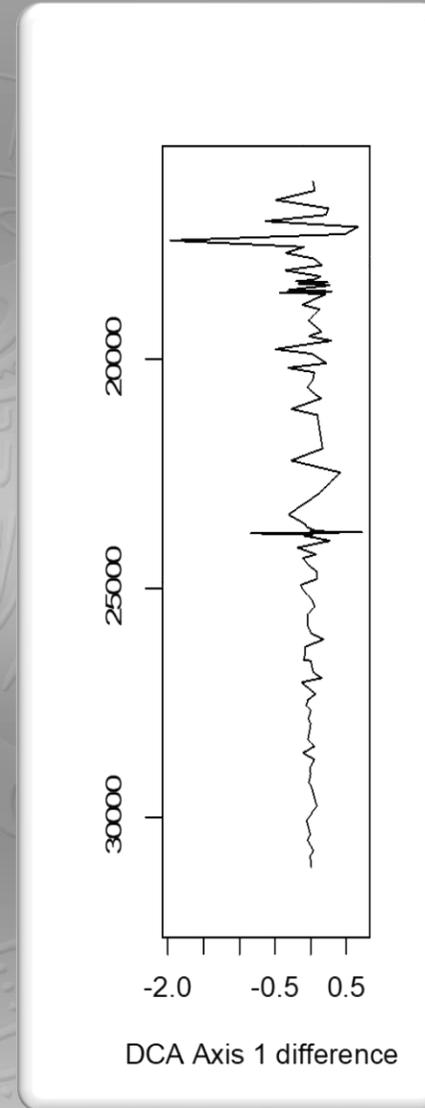
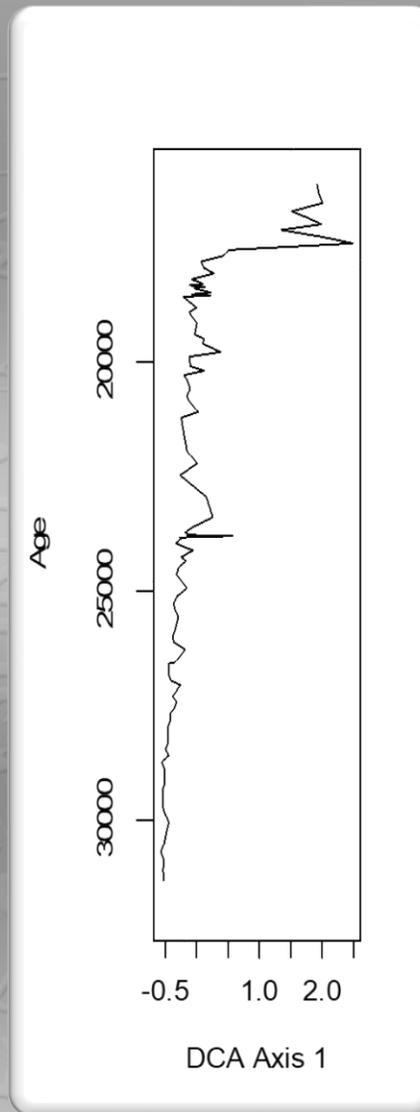


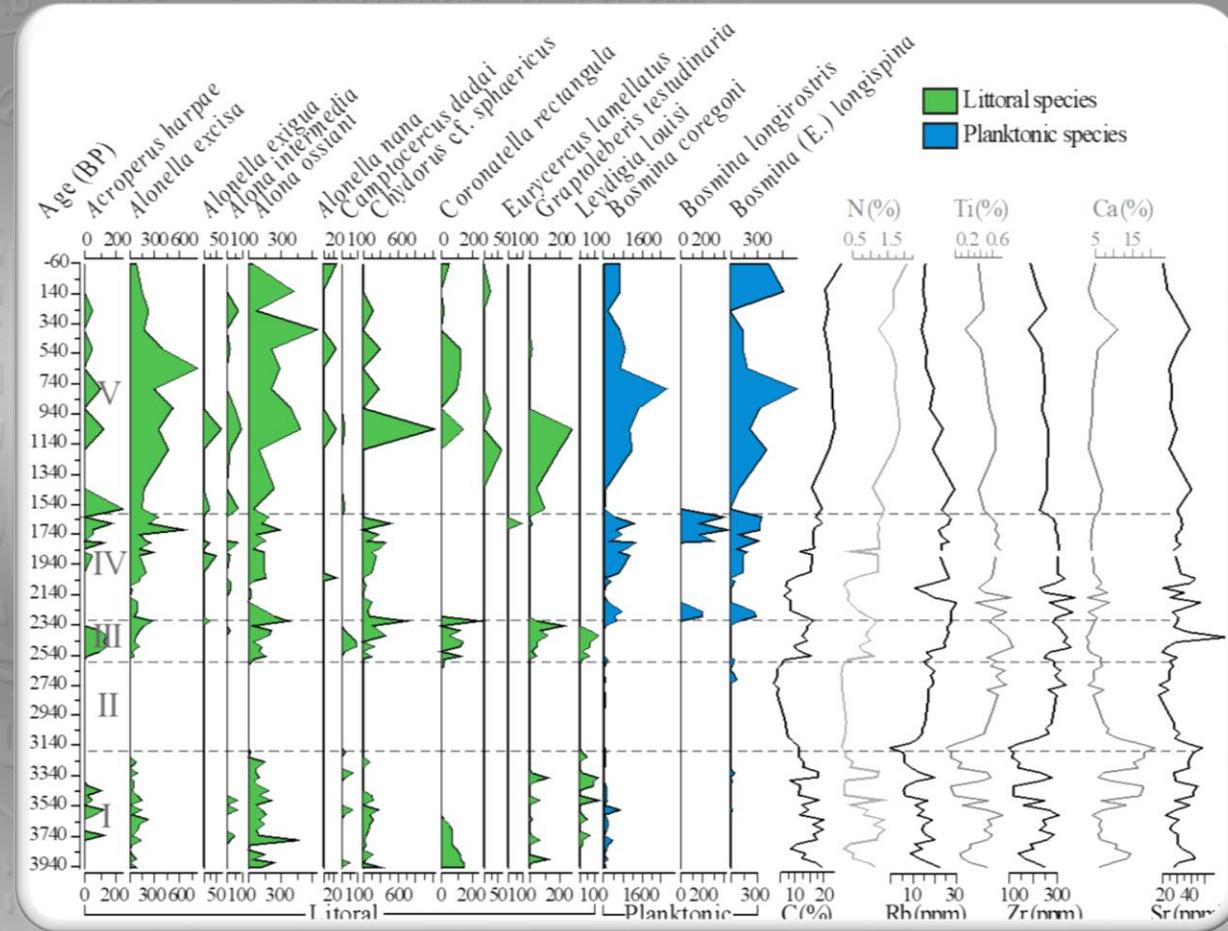
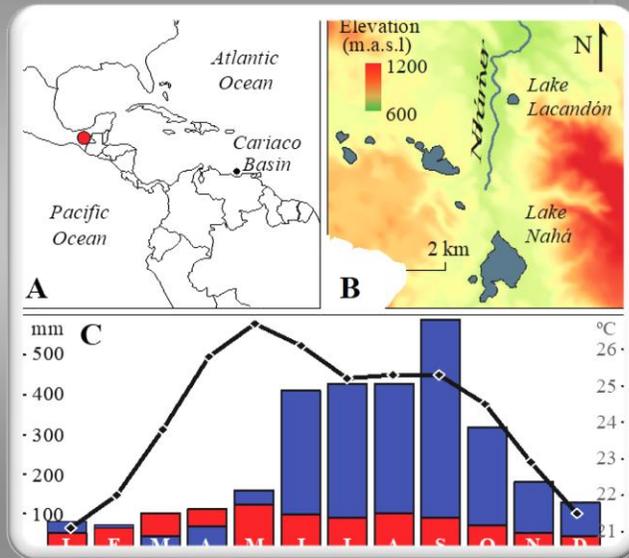
- Una vez identificados los gradientes abientales asociados con los ejes, estos pueden ser graficados estratigráficamente para facilidad de interpretación

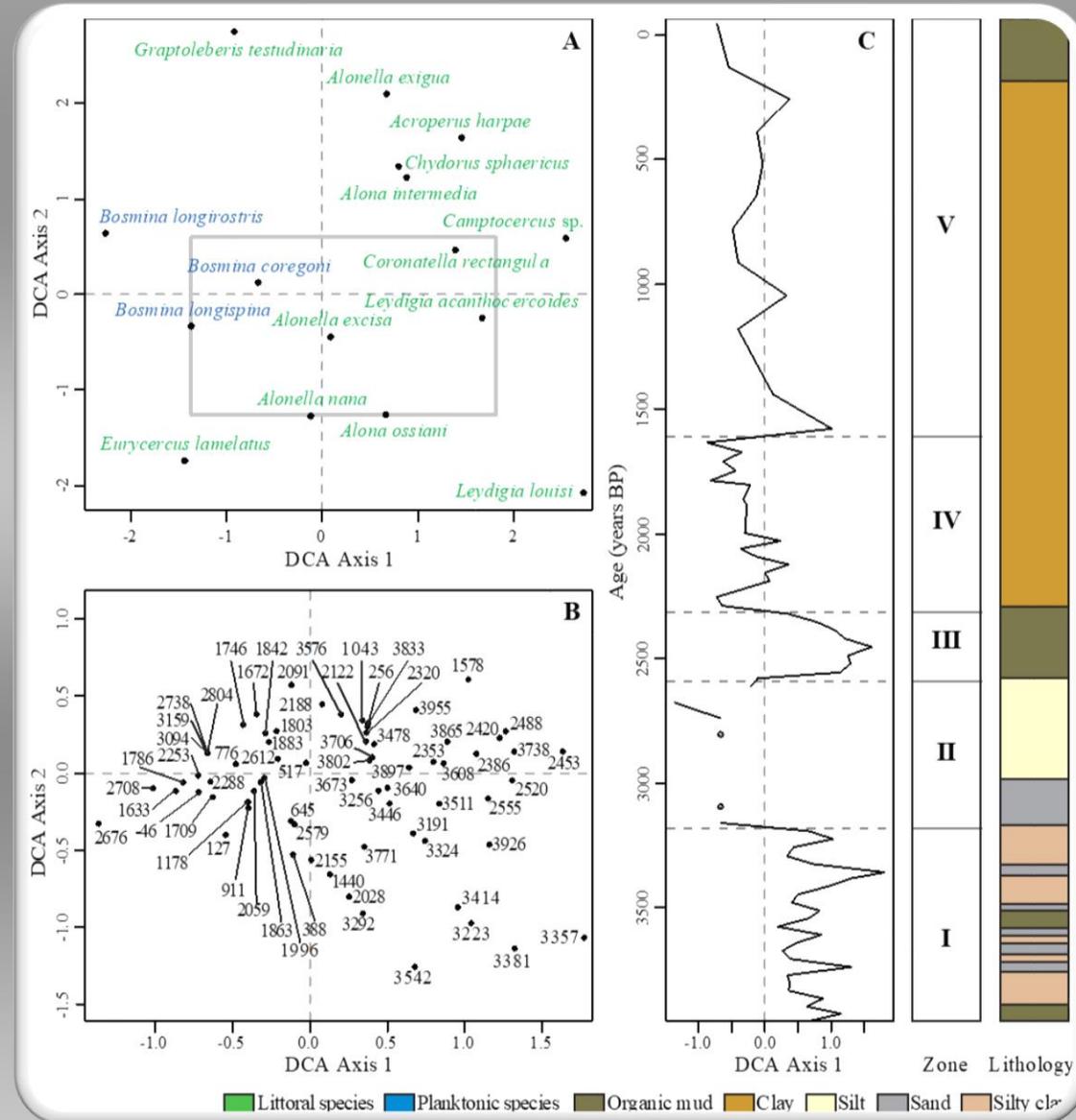
```
plot(scores.samples[,1],pi6.metadata[2],t="I",ylab="Age",xlab="DCA Axis 1",ylim=c(32000,16000))
```

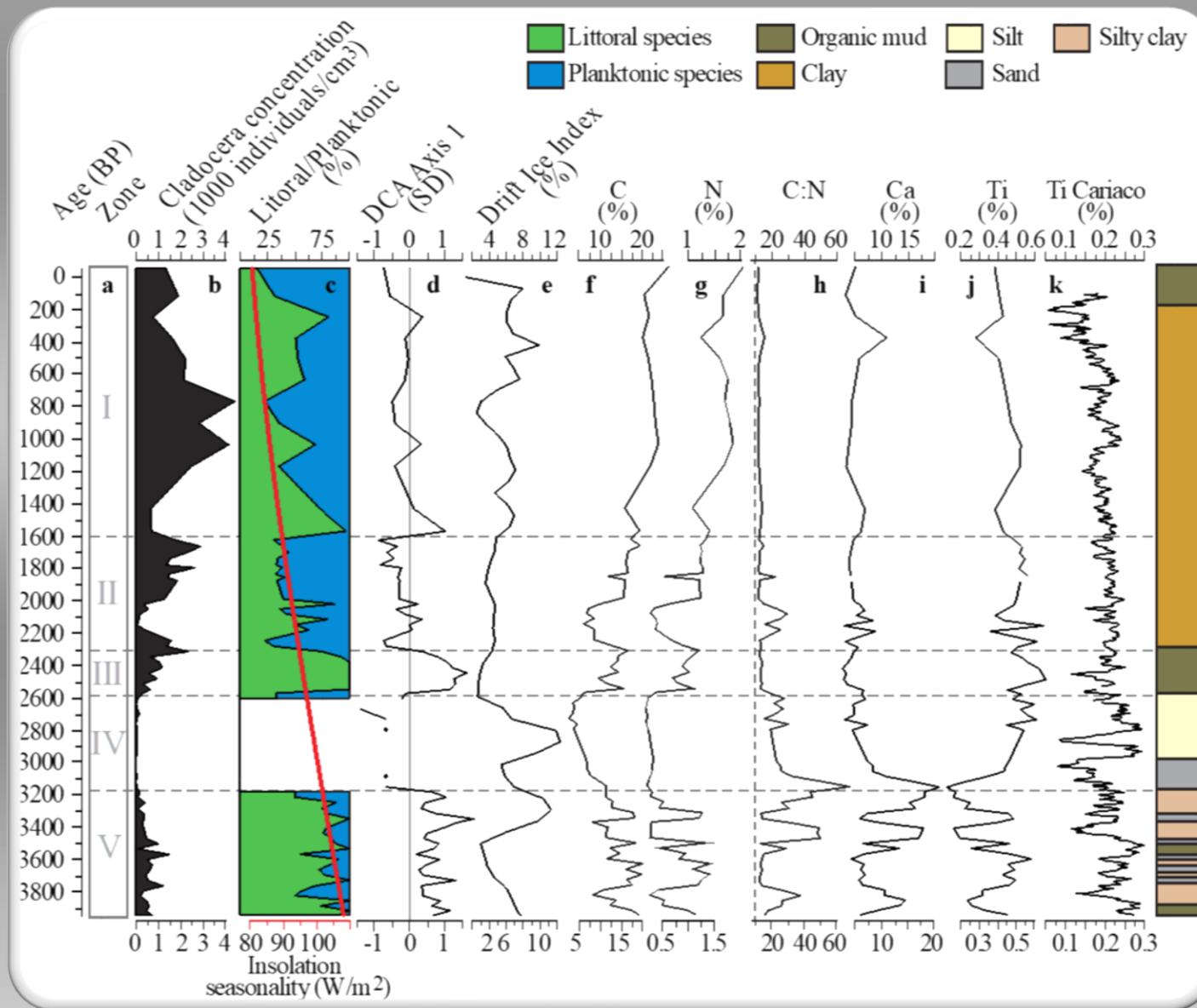
- Dada la metodología utilizada, los ejes tienen como unidades desviaciones estándar, mismas que pueden ser interpretadas en términos de recambio ecológico.

```
plot(scores.samples[-1,1]-scores.samples[-100,1],pi6.metadata[-100,2],t="I",ylab="",xlab="DCA Axis 1 difference",ylim=c(32000,16000))
```

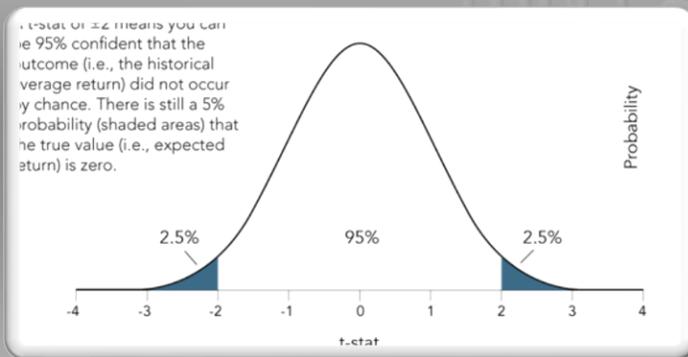






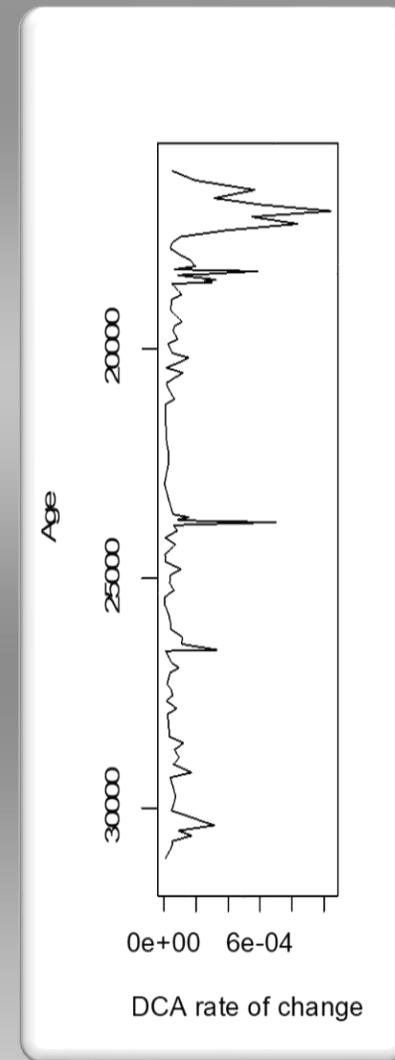
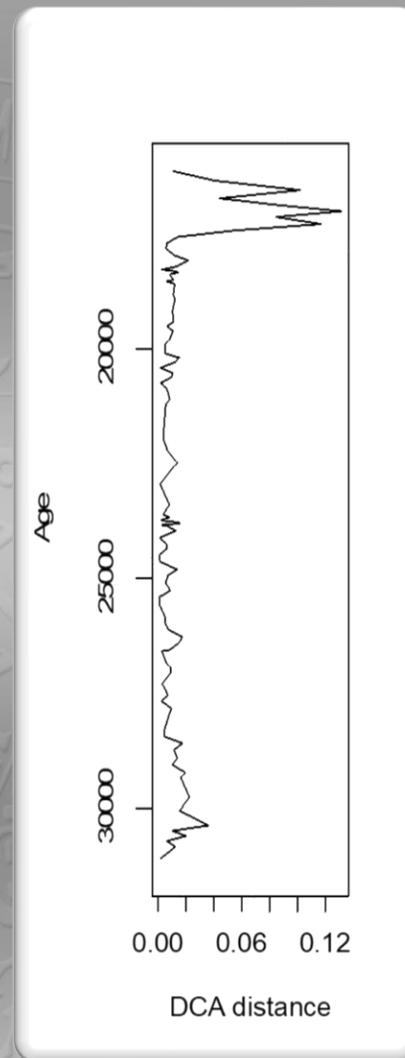
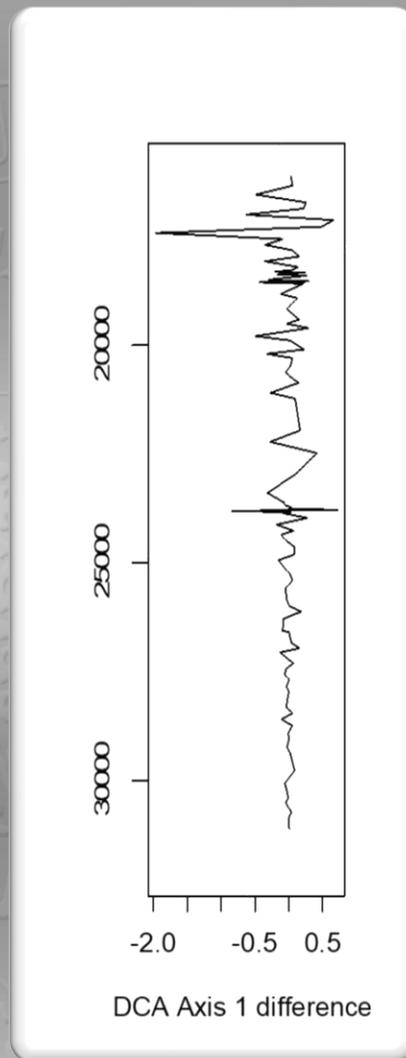


- Todos los ejes tiene como unidades desviaciones estándar, de manera que el cálculo de distancia entre muestras en el nuevo espacio coordinado proporciona una aproximación a la distancia ecológica.
- Las distancias ecológicas se pueden interpretar como cuantiles de una distribución de t con grados de libertad muestras-especies, y recambio ecológico equivalente a la probabilidad asociada.



```
change(pi6.pollenp[,3:146],pi6.metadata[,2],dca=T)
```

```
change(pi6.pollenp[,3:146],pi6.metadata[,2],dca=T,roc=T)
```



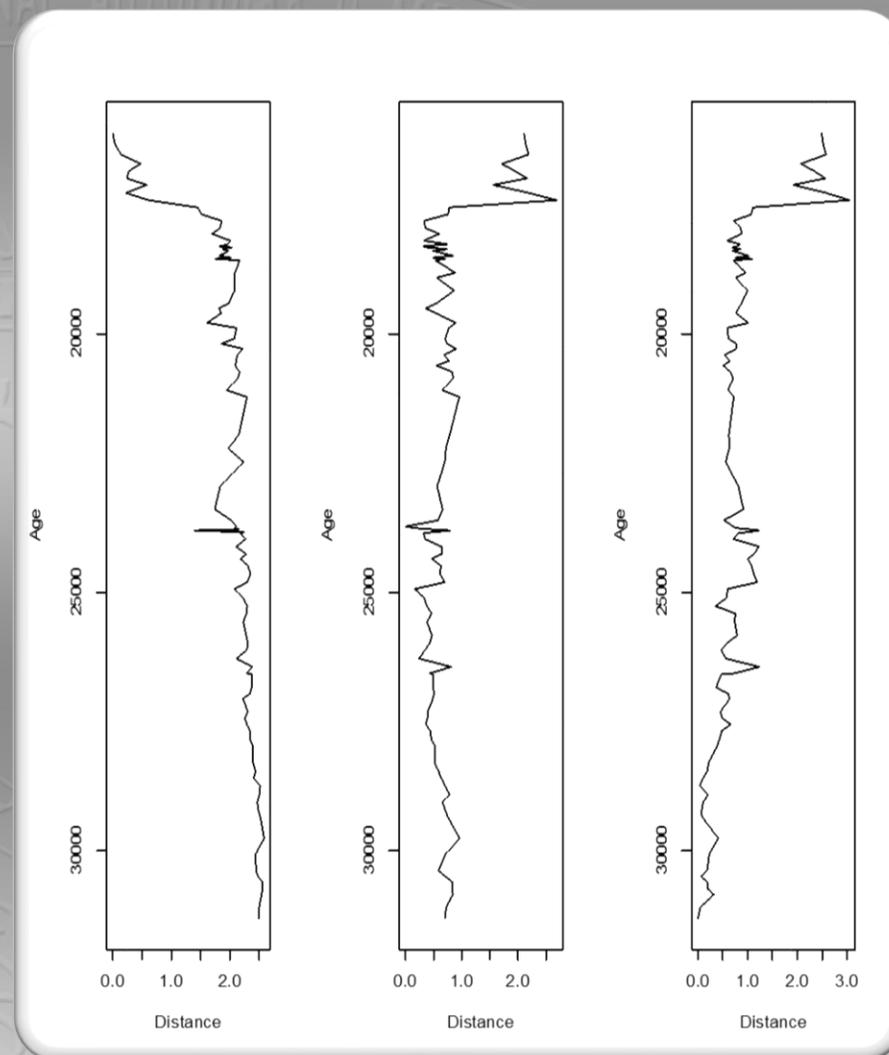
- Al igual que entre muestras contiguas, las distancias se pueden calcular con respecto a una muestra en particular.

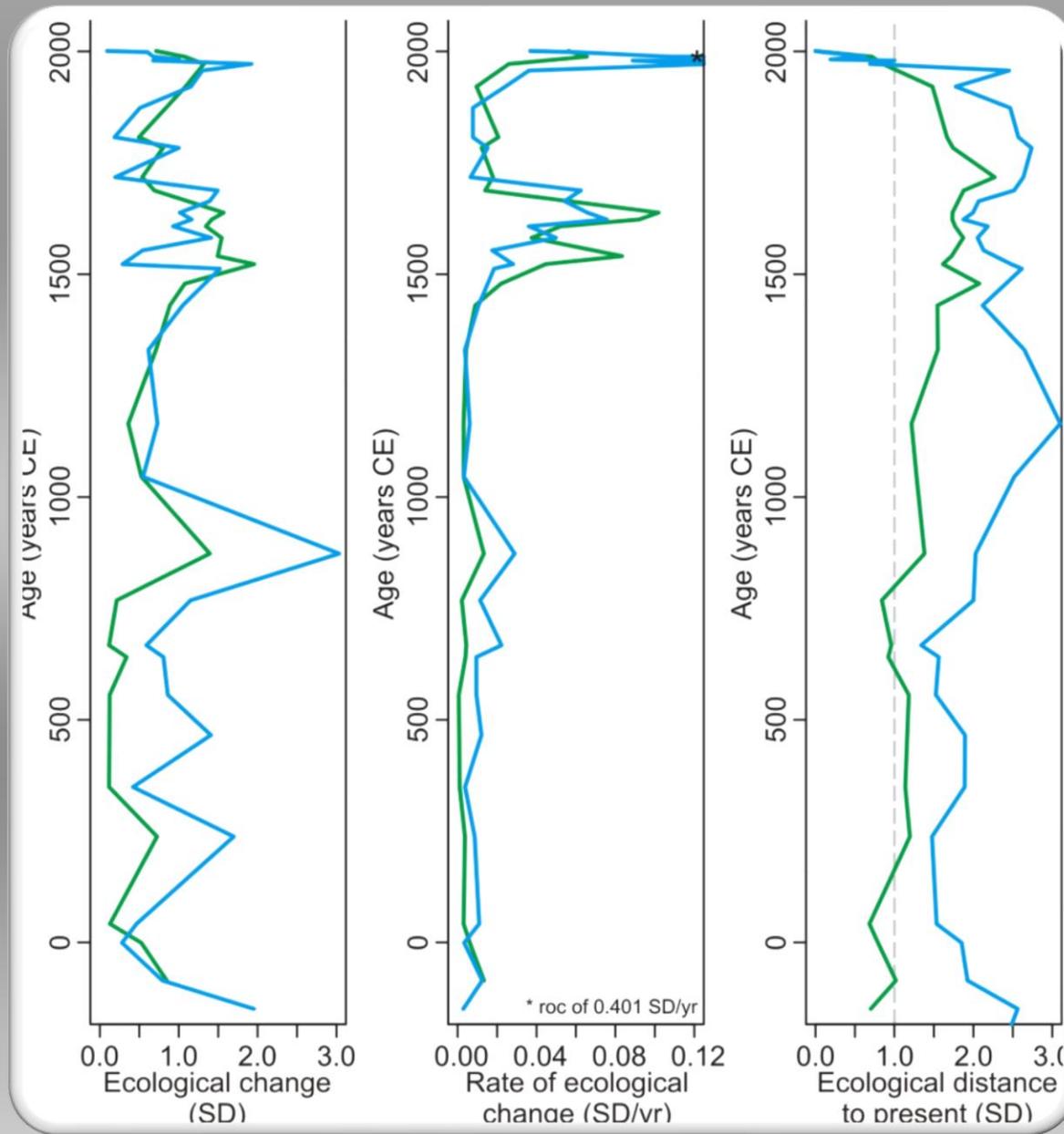
```
par(mfrow=c(1,3))
```

```
analog.sing(pi6.pollenp[,3:146],base=1,pi6.metadata[2],dca=T)
```

```
analog.sing(pi6.pollenp[,3:146],base=50,pi6.metadata[2],dca=T)
```

```
analog.sing(pi6.pollenp[,3:146],base=100,pi6.metadata[2],dca=T)
```



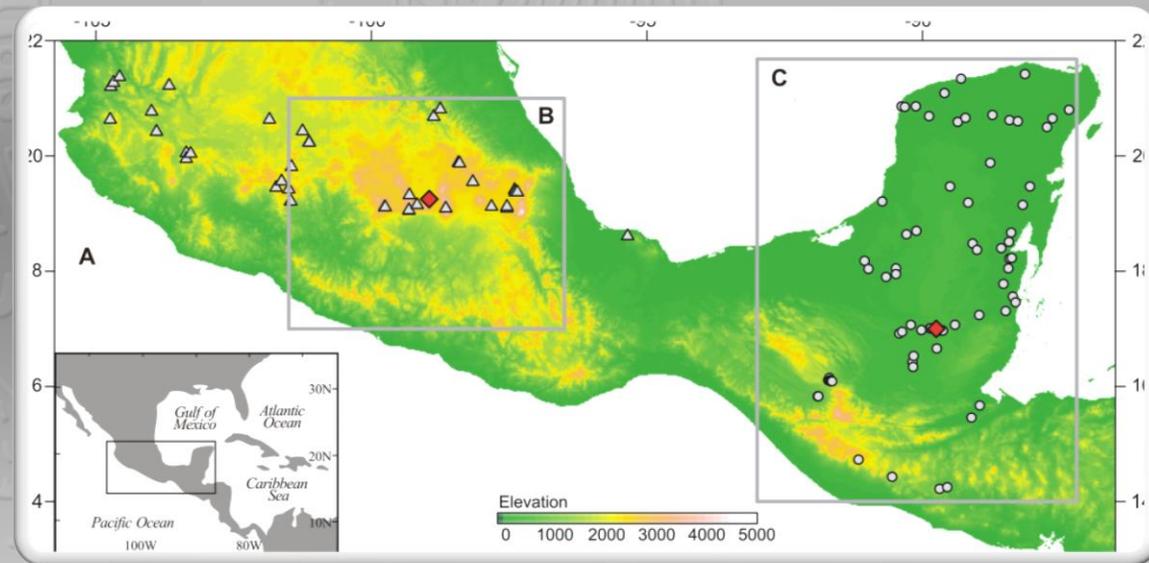


Cuantificación de cambios ambientales y ecológicos: Interpretación *a posteriori*

- Reconstrucción ambiental y ecológica basada en datos paleo en conjunción con datos modernos.
- La comparación entre datos modernos y datos fósiles permite la reconstrucción del pasado.
- Los datos modernos representan el espacio ecológico y ambiental posible mediante dos aproximaciones: i) la realidad literal representada en los datos (e.g. MAT, WA), y ii) la realidad modelada a partir de los datos modernos (WA-PLS).

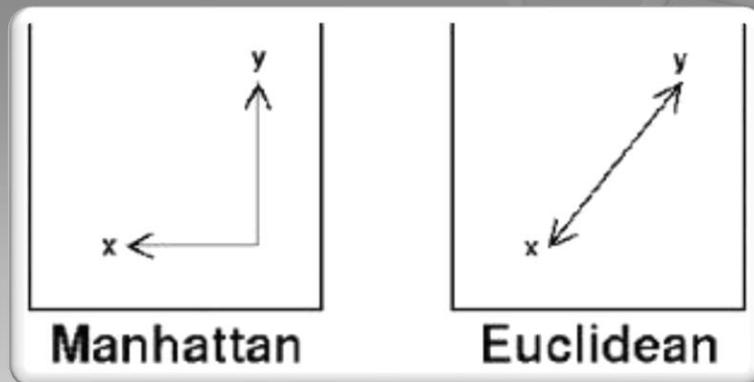
```
load("modern.pollenp.RData")
```

```
apply(modern.pollenp[,1:2],2,range)
```



Cuantificación de cambios ambientales y ecológicos: lo moderno como analogía del pasado

- Coincidencias entre el pasado y el presente permiten la identificación de los patrones temporales del pasado sobre las estructuras espaciales del presente.
- Índices de similitud y distancia entre lo moderno y lo fósil permiten una aproximación a las condiciones del pasado.
- Los resultados dependen del tipo de índice que se utilice para la comparación (e.g. distancia Euclídiana, distancia de Manhattan).



```
colnames(modern.pollenp)
```

```
colnames(pi6.pollenp)
```

```
intersect(colnames(modern.pollenp),colnames(pi6.pollenp))
```

```
dist(rbind(modern.pollenp[,common],pi6.pollenp[,common]))->dist.euc
```

```
as.matrix(dist.euc)->dist.euc
```

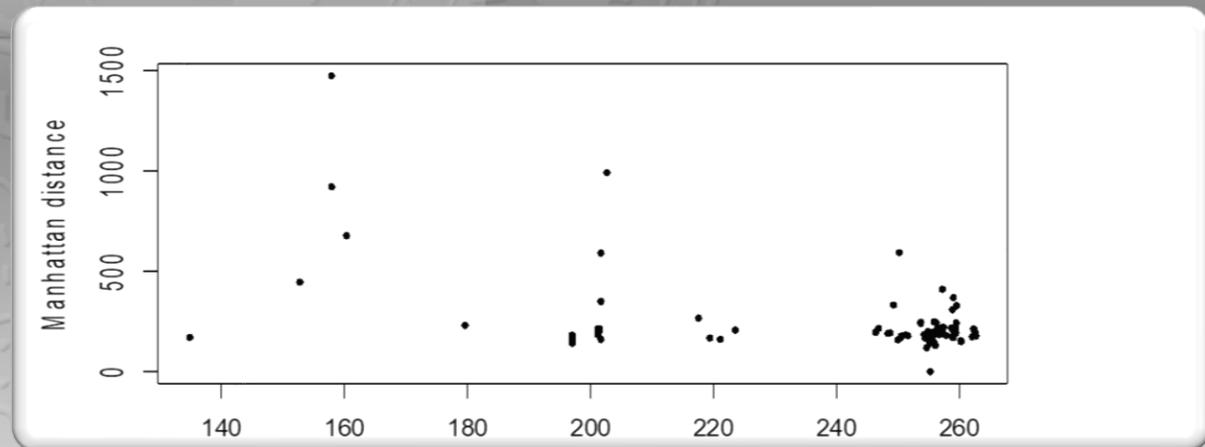
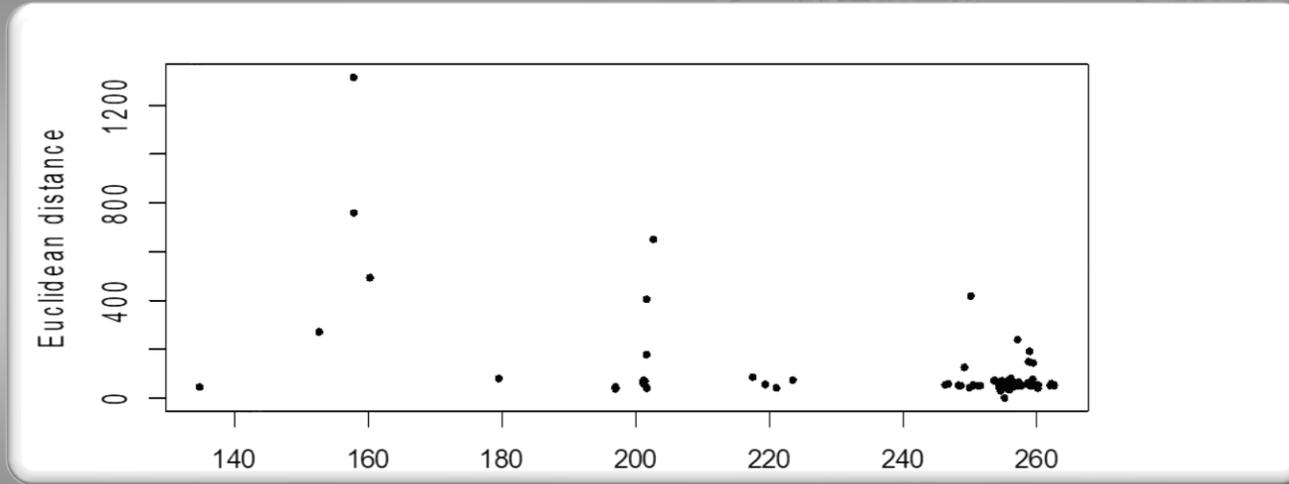
```
dist(rbind(modern.pollenp[,common],pi6.pollenp[,common]), method="manhattan")->dist.man
```

```
as.matrix(dist.man)->dist.man
```

La distancia a lo moderno (16125 BP)

```
plot(modern.pollenp[,3],dist.euc[1:81,1],pch=16,cex=0.7,xlab="Temperature",ylab="Euclidean distance")
```

```
plot(modern.pollenp[,3],dist.man[1:81,1],pch=16,cex=0.7,xlab="Temperature",ylab="Manhattan distance")
```



Los análogos modernos en la reconstrucción paleoambiental

- Paquete rioja ofrece herramientas para la reconstrucción ambiental.

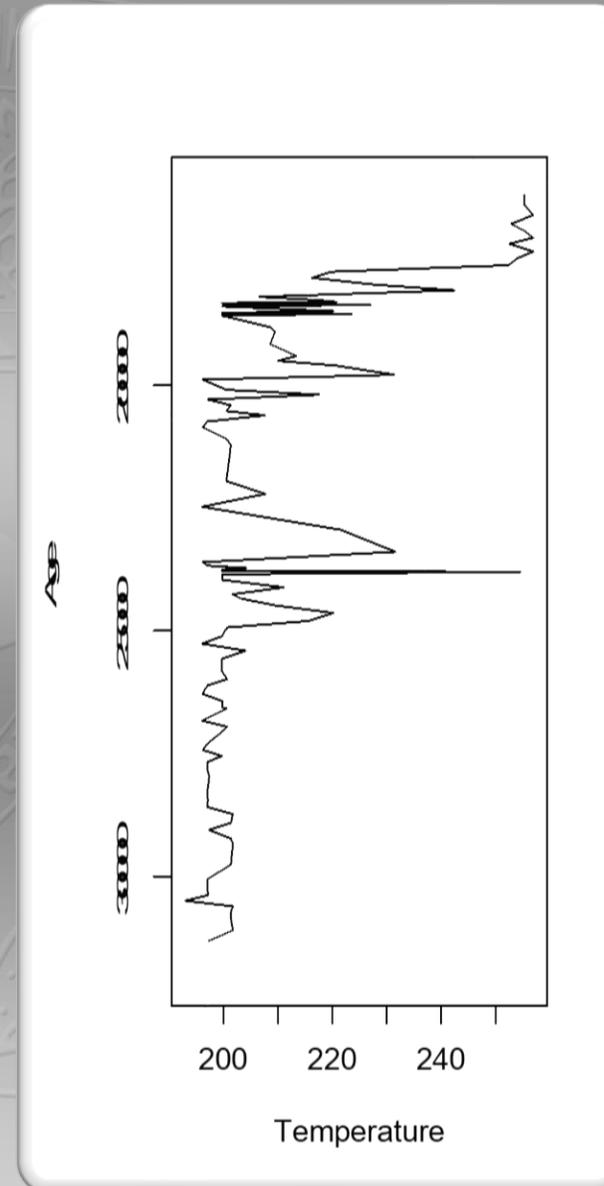
```
install.packages("rioja",dependencies)
```

```
library(rioja)
```

```
MAT(modern.pollenp[,4:152],modern.pollenp[,3])-  
>temp.mat
```

```
predict(temp.mat,pi6.pollenp[,3:146])->mat.recons
```

```
plot(mat.recons$fit[,1],pi6.metadata[,2],t="l",ylab="Age",  
xlab="Temperature",ylim=c(32000,16000))
```



WA-PLS: Lo moderno usado para generalizar relaciones especies-clima

- Mezcla entre medias ponderadas (los porcentajes de las especies pesan sobre el promedio de las condiciones ambientales).

```
WAPLS(modern.pollenp[,4:152],modern.pollenp[,3])->  
temp.wapls
```

```
predict(temp.wapls,pi6.pollenp[,3:146])->wapls.recons
```

```
plot(wapls.recons$fit[,1],pi6.metadata[,2],t="l",ylab="Age",  
xlab="Temperature",ylim=c(32000,16000))
```

```
lines(wapls.recons$fit[,2],pi6.metadata[,2],col="red")
```

```
lines(wapls.recons$fit[,3],pi6.metadata[,2],col="blue")
```

```
lines(wapls.recons$fit[,4],pi6.metadata[,2],col="orange")
```

