

Trabajo practico 1

Ejercicio 5

Grupo

12 mayo, 2021

A) Halle la distancia genética de Prevosti entre variedades

Para calcular esta distancia usamos la siguiente sentencia

```
matriz_distancia <- dist(datos_df, method = "manhattan", diag = TRUE) / 33
```

Mostramos la matriz resultante de manera parcial, ya que es muy grande como para mostrarla de manera completa en una pagina.

	A1	A2	A3	A4	A5	A6	A7	A8	F1	F2
A1	0.00	0.27	0.30	0.33	0.30	0.39	0.39	0.33	0.55	0.47
A2	0.27	0.00	0.33	0.30	0.33	0.30	0.30	0.24	0.45	0.33
A3	0.30	0.33	0.00	0.03	0.18	0.15	0.15	0.09	0.42	0.41
A4	0.33	0.30	0.03	0.00	0.15	0.12	0.12	0.06	0.39	0.38
A5	0.30	0.33	0.18	0.15	0.00	0.09	0.09	0.15	0.42	0.41
A6	0.39	0.30	0.15	0.12	0.09	0.00	0.00	0.06	0.39	0.38
A7	0.39	0.30	0.15	0.12	0.09	0.00	0.00	0.06	0.39	0.38
A8	0.33	0.24	0.09	0.06	0.15	0.06	0.06	0.00	0.45	0.38
F1	0.55	0.45	0.42	0.39	0.42	0.39	0.39	0.45	0.00	0.18
F2	0.47	0.33	0.41	0.38	0.41	0.38	0.38	0.38	0.18	0.00

B) Podría aplicar el coeficiente de similitud SM ? Porque ?

Si, pero no lo hacemos porque perderiamos informacion, ya que las bandas presentan mas de 2 valores posibles. Categorizar los valores observados en solamente dos categorias implicaria una perdida de informacion.

C) Realice un Análisis de Coordenadas Principales para encontrar la configuración de las variedades de pepino en función de esta caracterización molecular. Encuentra asociaciones en función del tipo de pepino?

En la Figura 1 podemos ver que los pepinos del tipo F suelen encontrarse en el primer cuadrante, los del tipo P en el segundo cuadrante, los de tipo A en el tercero, y los de tipo H en el cuarto. Sin embargo esta ordenacion es imprecisa, ya que por ejemplo, hay pepinos de los tipo F y A en el segundo cuadrante. Sino estuvieran los colores que indican los tipos de pepinos, probablemente obtendriamos agrupamientos que estuvieran compuestos en su mayoria un unico tipo de pepino, pero que tambien incluirian pepinos de otros tipos.

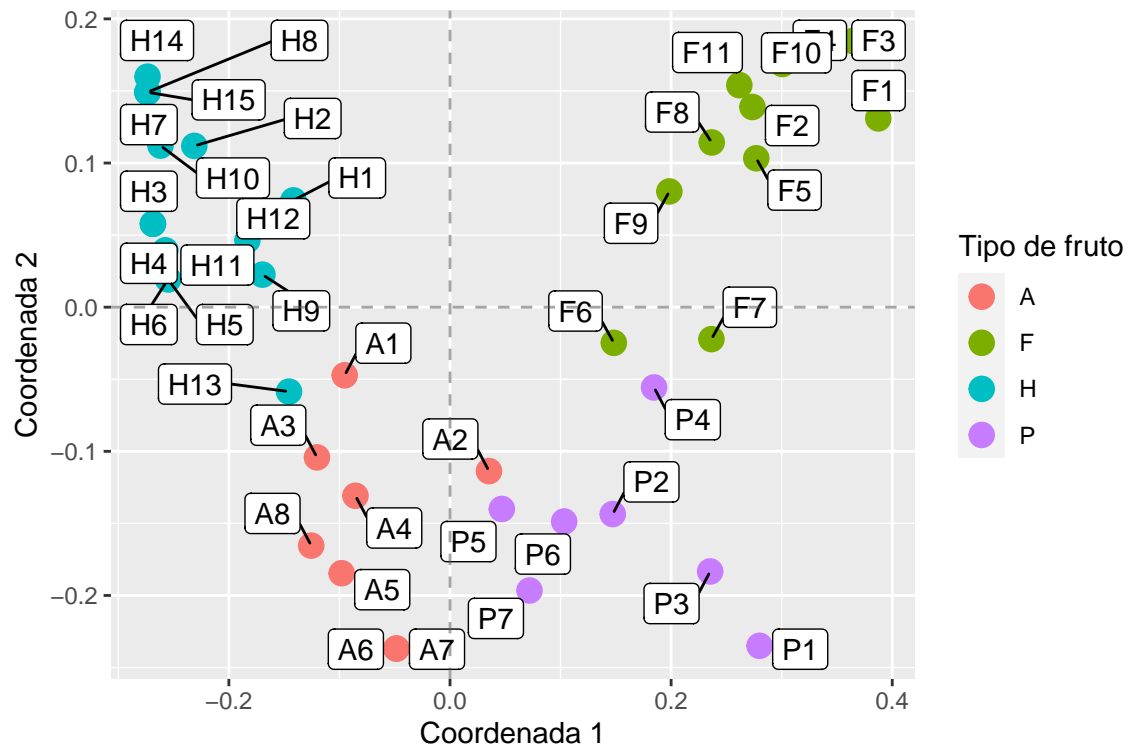


Figure 1: Caracterización molecular de las variedades de pepino.

D) Encuentre el dendrograma ultramétrico con ligamiento UPGMA

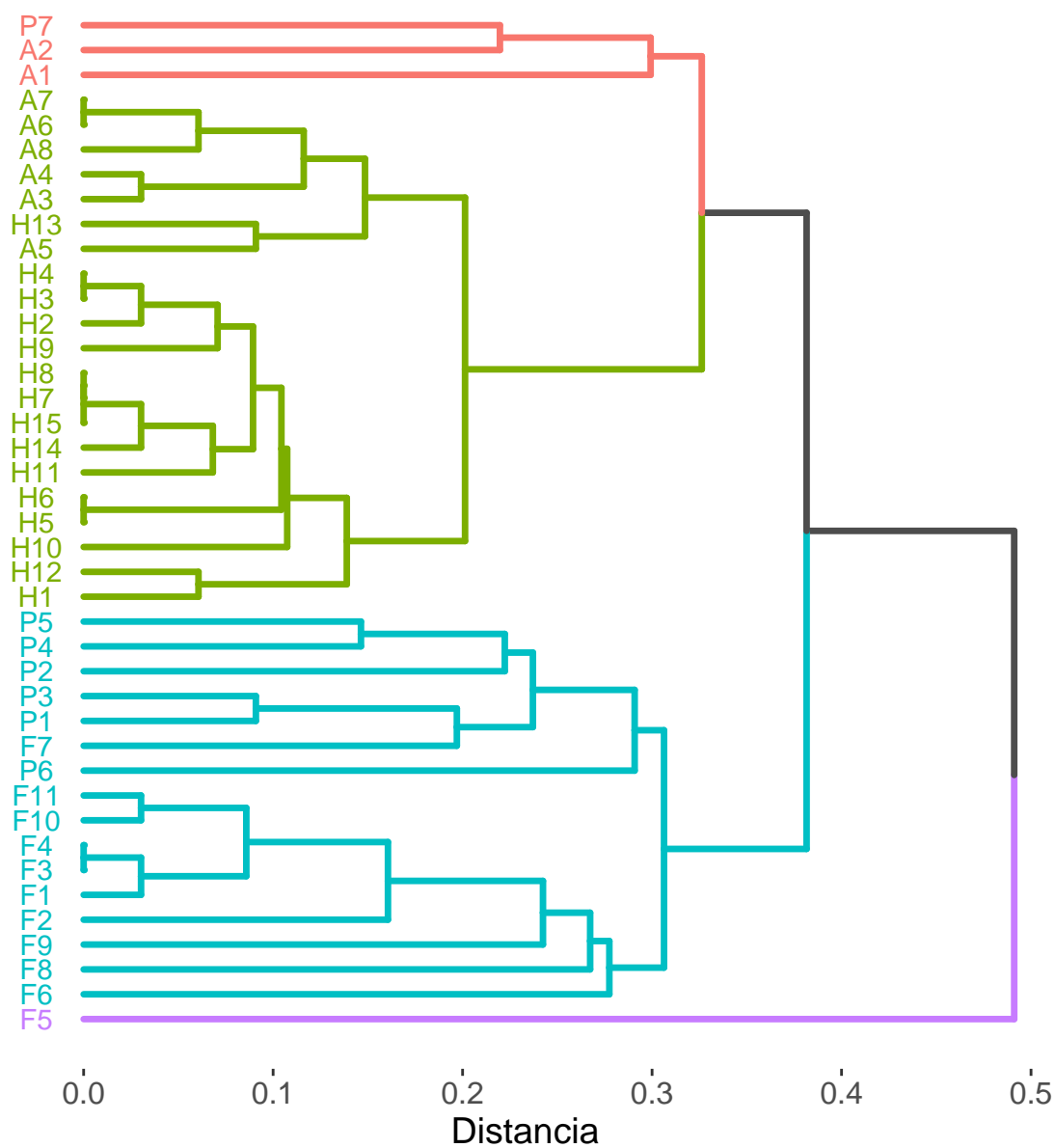


Figure 2: Dendrograma Ultrametrico con ligamiento UPGMA.

E) Mida a través de su matriz cofenética la concordancia con la matriz de distancias que le dio origen

La concordancia entre la matriz de distancias cofenética y la matriz de distancia original es igual a 0.872. En la Figura 3 se puede ver la asociación positiva entre las dos medidas de distancia. La dispersión en la nube de puntos aumenta a medida que la distancia es mayor, lo que significa que las dos medidas de distancia tienden a diferir más cuando la distancia entre los tipos de pepinos es mayor.

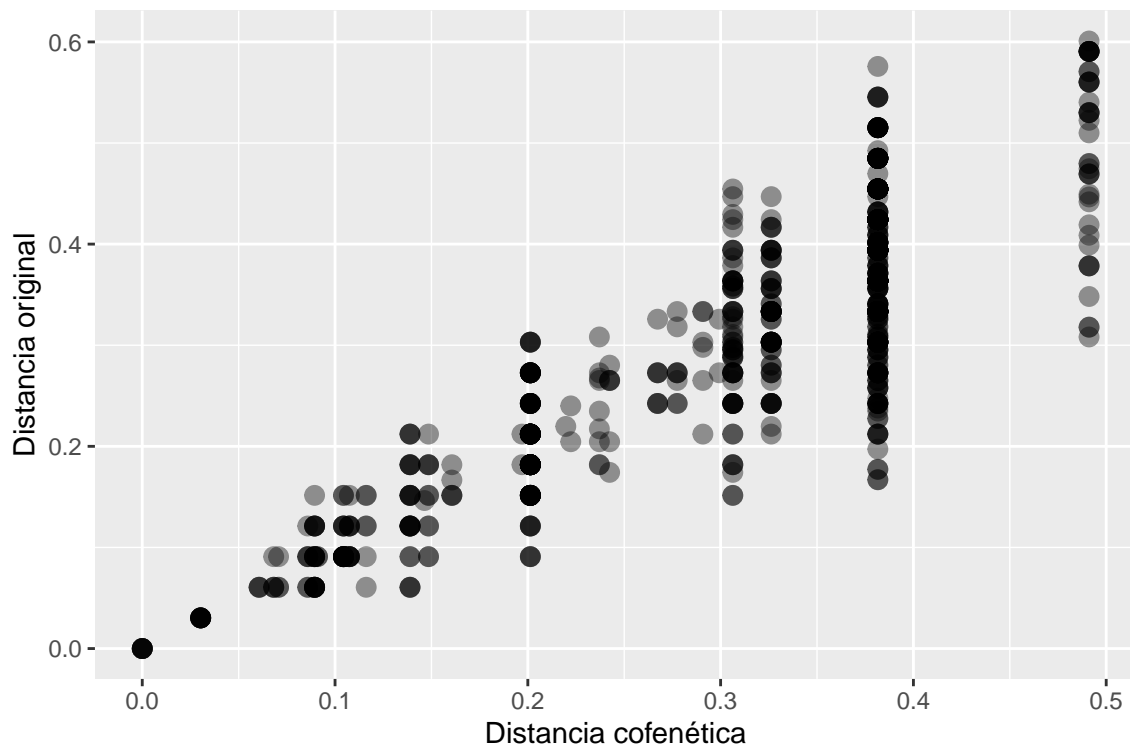
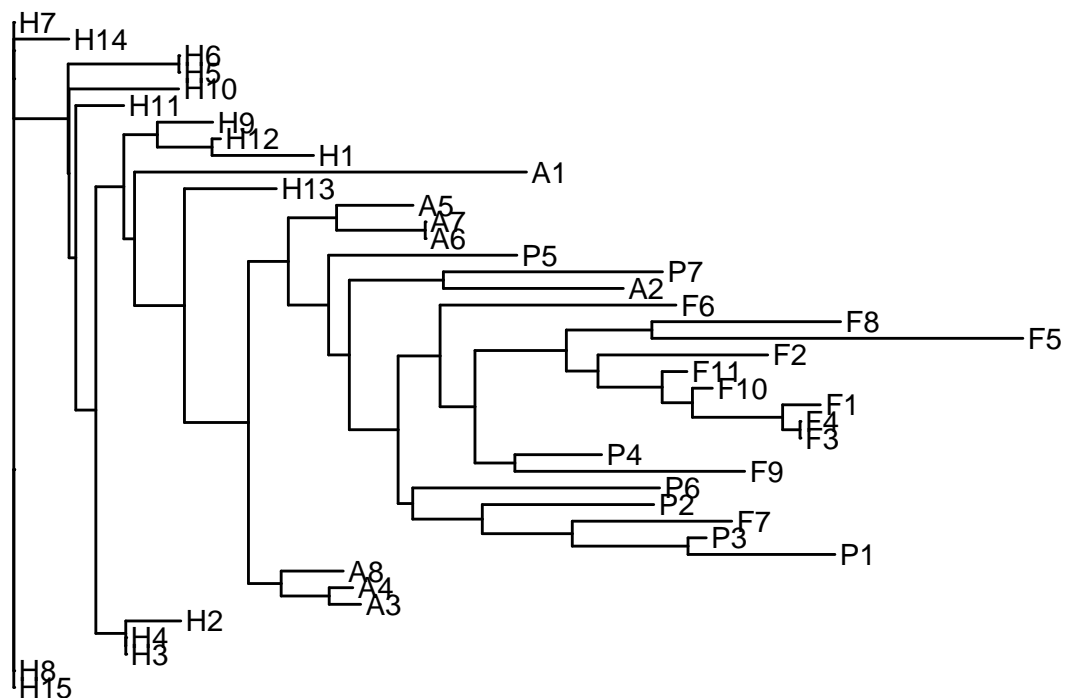


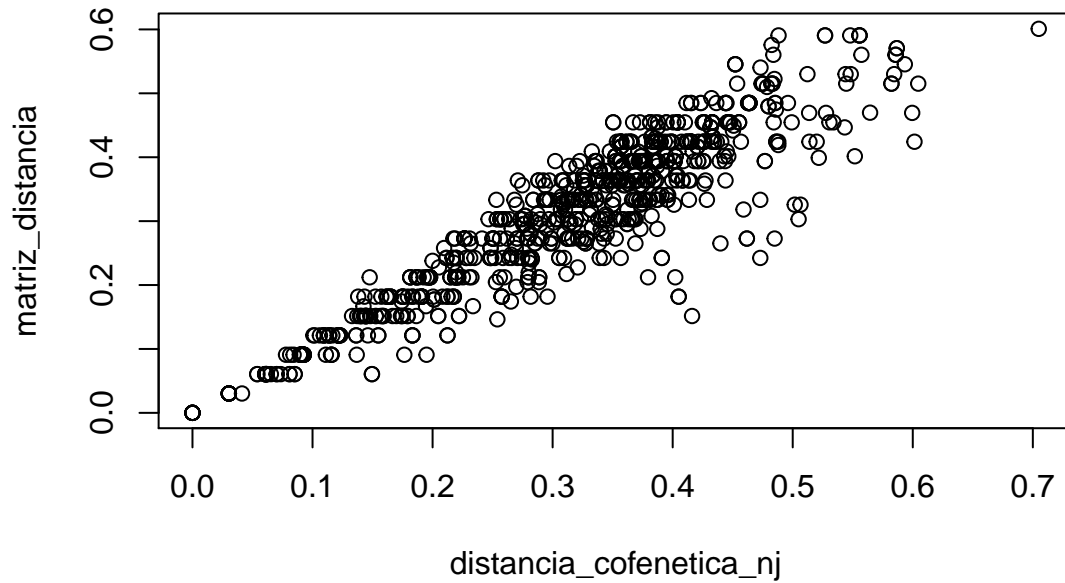
Figure 3: Grafico de dispersion entre distancia original y distancia cofenética.

F) Halle el dendograma aditivo Neighbor Joining



G) Mida su concordancia con matriz de distancia original

La concordancia entre la matriz de distancias cofenética construida a partir del dendograma aditivo Neighbor Joining y la matriz de distancia original es igual a 0.926. **falta interpretar**



H) Relacione ambos dendogramas y saque conclusiones

Falta sacar conclusiones.

La concordancia es 0.872. **interpretar**