

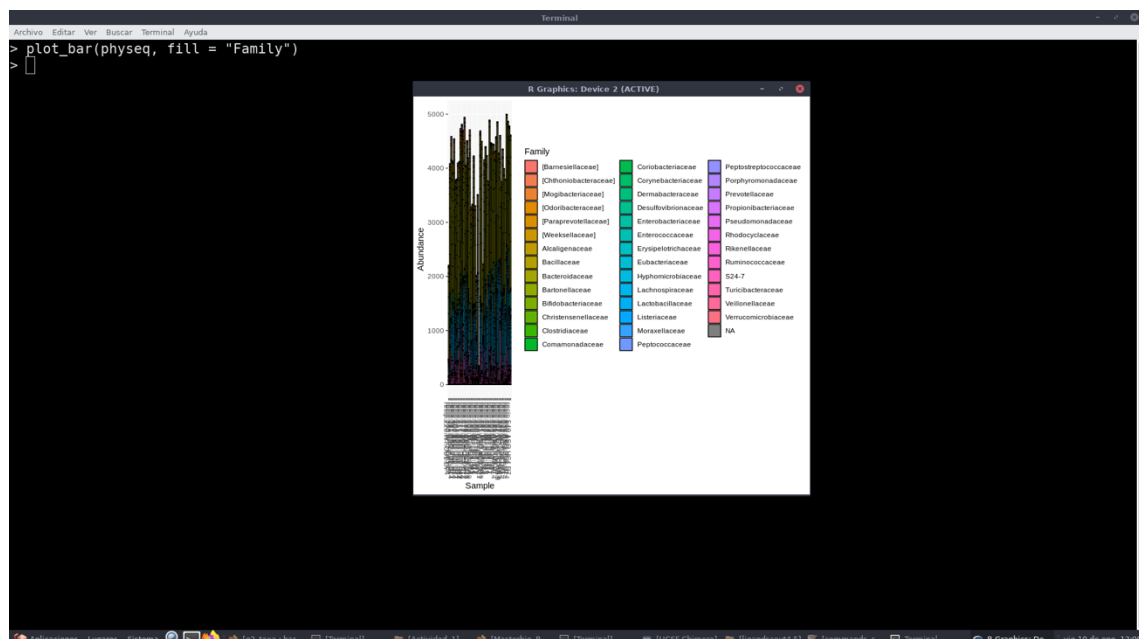
1. Haz un subset de la orden Clostridiales, ¿Cuántos taxones había y cuántos taxones hay ahora?

```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
> physeq.clos = subset_taxa(physeq, Order == "Clostridiales")
> physeq.clos
phyloseq-class experiment-level object
otu_table() OTU Table: [ 156 taxa and 48 samples ]
sample_data() Sample Data: [ 48 samples by 9 sample variables ]
tax_table() Taxonomy Table: [ 156 taxa by 7 taxonomic ranks ]
phy_tree() Phylogenetic Tree: [ 156 tips and 155 internal nodes ]
> physeq
phyloseq-class experiment-level object
otu_table() OTU Table: [ 287 taxa and 48 samples ]
sample_data() Sample Data: [ 48 samples by 9 sample variables ]
tax_table() Taxonomy Table: [ 287 taxa by 7 taxonomic ranks ]
phy_tree() Phylogenetic Tree: [ 287 tips and 286 internal nodes ]
>
```

Al realizar el subset del orden Clostridiales con el código: `physeq.clos = subset_taxa(physeq, Order == "Clostridiales")`, antes había 287 taxones, mientras que ahora quedan 156 taxones, al filtrar por el orden Clostridiales.

2. Representa para cada muestra su abundancia para cada una de las familias.

Con la función `plot_bar(physeq, fill = "Family")`, filtrando por el campo family obtenemos para cada muestra su abundancia para cada una de las familias.



Y calculamos el valor medio del índice de Shannon para cada grupo:

```
mean(spl$Healthy)
```

```
mean(spl$PD)
```

- Después se realiza el test de Kolmogorov-Smirnov para obtener el p-valor, y se realiza un ajuste del p-valor obtenido.

```
pv <- ks.test(spl$Healthy, spl$PD)$p.value
```

```
padj <- p.adjust(pv)
```

```
Terminal
> d = meta(physeq)
> d$diversity <- microbiome::diversity(physeq, "shannon")$shannon
> d$diversity
[1] 2.369679 2.689382 2.863240 3.041673 3.000449 2.972395 2.842200 3.046244
[9] 2.642987 3.057670 2.916675 3.117208 3.114286 2.715667 2.892660 3.380230
[17] 2.965027 2.679207 2.581345 2.876582 2.561989 3.472265 2.563211 2.268194
[25] 2.882049 2.530586 2.517621 2.703146 2.634748 2.622611 2.027248 2.708270
[33] 2.944224 1.996112 3.037790 2.578882 2.375542 2.897508 1.972140 2.870625
[41] 2.597007 2.790595 2.235320 2.453876 2.404599 2.863774 2.603077 2.821195
> spl = split(d$diversity, d$donor_status)
> spl
$Healthy
[1] 2.369679 2.689382 2.863240 3.041673 3.000449 2.972395 2.842200 3.046244
[9] 2.642987 3.057670 2.916675 3.117208 3.114286 2.715667 2.892660 3.380230
[17] 2.965027 2.679207 2.581345 2.876582 2.561989 3.472265 2.882049 2.703146 2.870625

$PD
[1] 2.561989 2.563211 2.268194 2.530586 2.517621 2.634748 2.622611 2.027248
[9] 2.708270 2.944224 1.996112 3.037790 2.578882 2.375542 2.897508 1.972140
[17] 2.597007 2.790595 2.235320 2.453876 2.404599 2.863774 2.603077 2.821195

> mean(spl$Healthy)
[1] 2.90387
> mean(spl$PD)
[1] 2.541922
> pv <- ks.test(spl$Healthy, spl$PD)$p.value
> pv
[1] 0.0001040143
> padj <- p.adjust(pv)
> padj
[1] 0.0001040143
> █
```

4. Utiliza el paquete microbiome para obtener los índices de diversidad alfa, ¿Cuál es la muestra con menor índice de Shannon y qué valor tiene?

La muestra con menor índice de Shannon es: recip.466.ASO.PD3.D7, con un índice de Shannon de 1.972140.

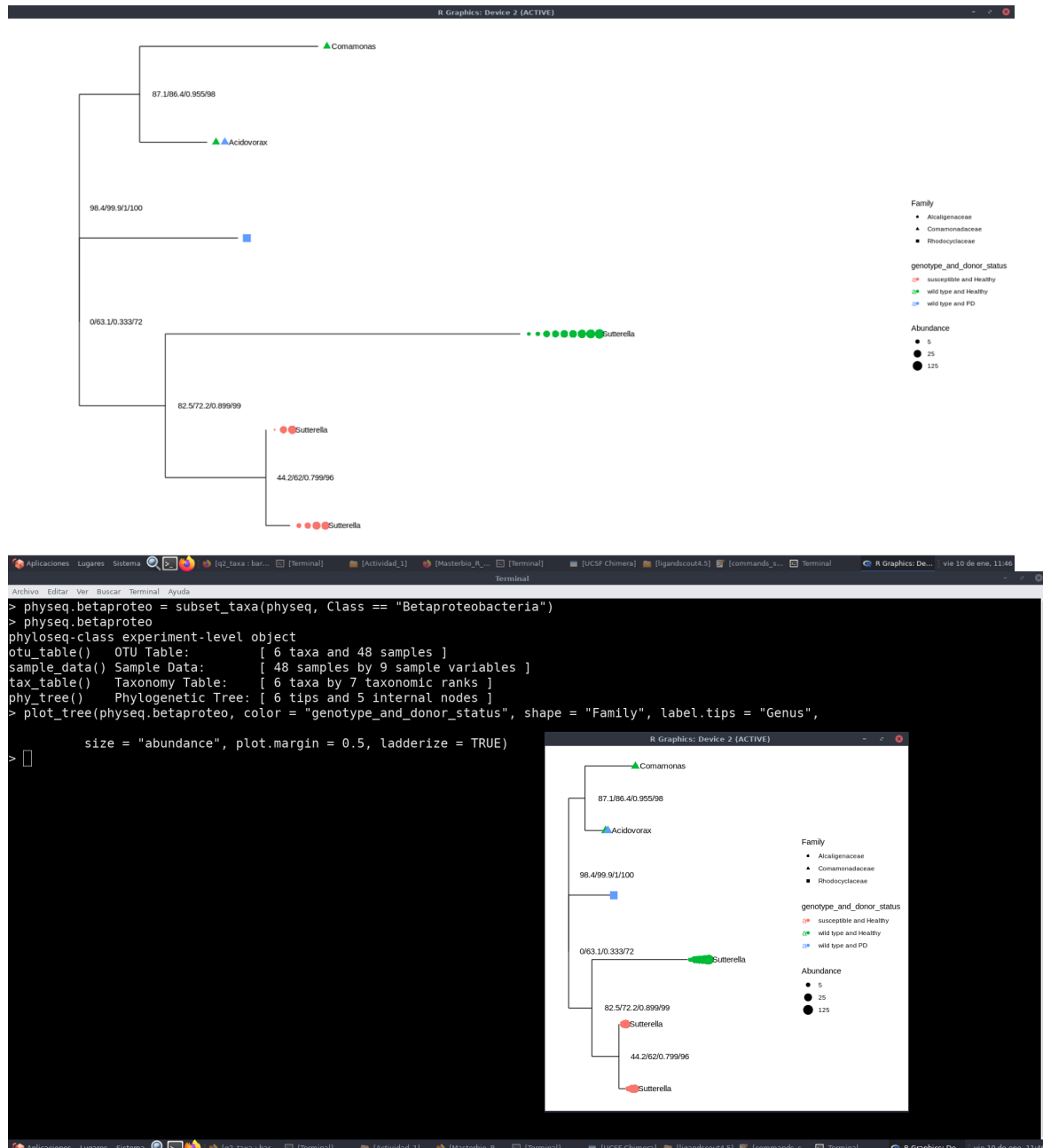
```
Terminal
> orden_shannon <- microbiome::diversity(ASV)[order(as.numeric(as.character(microbiome::diversity(ASV)$shannon)), decreasing = FALSE),]
> View(orden_shannon)
> head(orden_shannon)
      inverse_simpson  gini_simpson  shannon  fisher coverage
recip.466.ASO.PD3.D7      4.703581      0.787396 1.972140  0.015096      2
recip.463.WT.PD3.D21      4.773133      0.790494 1.996112  4.323760      2
recip.462.WT.PD3.D49      4.490889      0.777327 2.027248  4.626499      2
recip.537.WT.PD4.D21      6.513110      0.846463 2.235320  4.405411      3
recip.460.WT.HC3.D49      7.876562      0.873041 2.268194  2.665856      3
recip.220.WT.OB1.D7       5.265387      0.810080 2.369679  6.109143      2
> █

Data: orden_shannon


| row.names                 | inverse_simpson | gini_simpson | shannon  | fisher   | coverage |
|---------------------------|-----------------|--------------|----------|----------|----------|
| 1 recip.466.ASO.PD3.D7    | 4.703581        | 0.787396     | 1.972140 | 0.015096 | 2        |
| 2 recip.463.WT.PD3.D21    | 4.773133        | 0.790494     | 1.996112 | 4.323760 | 2        |
| 3 recip.462.WT.PD3.D49    | 4.490889        | 0.777327     | 2.027248 | 4.626499 | 2        |
| 4 recip.537.WT.PD4.D21    | 6.513110        | 0.846463     | 2.235320 | 4.405411 | 3        |
| 5 recip.460.WT.HC3.D49    | 7.876562        | 0.873041     | 2.268194 | 2.665856 | 3        |
| 6 recip.220.WT.OB1.D7     | 5.265387        | 0.810080     | 2.369679 | 6.109143 | 2        |
| 7 recip.465.ASO.PD3.D4    | 6.754492        | 0.891392     | 2.375542 | 5.952913 | 3        |
| 8 recip.539.ASO.PD4.D4    | 6.986102        | 0.866927     | 2.445559 | 5.506201 | 3        |
| 9 recip.538.WT.PD4.D21    | 7.499091        | 0.866631     | 2.453076 | 6.423066 | 3        |
| 10 recip.461.ASO.HC3.D49  | 7.521543        | 0.867487     | 2.517621 | 6.099057 | 3        |
| 11 recip.461.ASO.HC3.D21  | 7.961439        | 0.874435     | 2.336568 | 5.344149 | 3        |
| 12 recip.459.WT.HC3.D49   | 6.895740        | 0.854932     | 2.561389 | 6.210832 | 3        |
| 13 recip.460.WT.HC3.D21   | 8.049380        | 0.875705     | 2.583211 | 6.320771 | 3        |
| 14 recip.464.WT.PD5.D4    | 6.251001        | 0.819802     | 2.519862 | 5.034983 | 3        |
| 15 recip.459.WT.HC3.D4    | 7.076974        | 0.898347     | 2.581346 | 6.392393 | 3        |
| 16 recip.467.WT.HC3.D49.D | 6.789581        | 0.855094     | 2.587007 | 6.623683 | 3        |
| 17 recip.540.ASO.HC4.D4   | 8.554095        | 0.884478     | 2.693077 | 6.595400 | 4        |
| 18 recip.462.WT.PD5.D21   | 8.944927        | 0.897036     | 2.622611 | 4.677102 | 4        |
| 19 recip.462.WT.PD5.D4    | 5.047495        | 0.889471     | 2.634748 | 5.364516 | 4        |
| 20 recip.405.ASO.PD3.D21  | 8.878385        | 0.888793     | 2.642987 | 5.248674 | 4        |
| 21 recip.459.ASO.HC3.D49  | 8.364940        | 0.880250     | 2.735007 | 5.546559 | 5        |
| 22 recip.290.ASO.D2C.D4   | 10.371236       | 0.939795     | 2.688882 | 5.435190 | 4        |
| 23 recip.461.ASO.HC3.D7   | 8.186262        | 0.877989     | 2.703146 | 5.774188 | 4        |
| 24 recip.462.WT.PD5.D7    | 8.359558        | 0.877453     | 2.719570 | 7.027489 | 5        |
| 25 recip.412.ASO.HC3.D7   | 8.878986        | 0.888794     | 2.719567 | 5.688510 | 4        |
| 26 recip.538.ASO.PD4.D49  | 3.627310        | 0.880251     | 2.789595 | 7.744636 | 4        |
| 27 recip.540.ASO.HC4.D7   | 10.426836       | 0.914104     | 2.621136 | 7.284832 | 4        |
| 28 recip.400.ASO.HC2.D4   | 12.210609       | 0.935447     | 2.842200 | 5.697006 | 5        |
| 29 recip.389.WT.HC2.D21   | 12.683260       | 0.917136     | 2.852440 | 6.027816 | 5        |


```

5. Haz un subset de la clase Betaproteobacteria y representa en un árbol filogenético cada uno de los géneros contenidos en esta clase, distinguiéndolos con un marcador que indique la familia en el árbol y coloreando las muestras según la variable "genotype_and_donor_status". Adjunta el código que utilizas para crearlo, así como el gráfico generado y qué conclusiones extraes de la observación de este árbol.



Código empleado:

Se hace un subset de la clase Betaproteobacteria

physeq.betaproteo = subset_taxa(physeq, Class == "Betaproteobacteria")

Y se hace el árbol filogenético coloreando por la variable "genotype_and_donor_status", usando una forma para distinguir cada familia y etiquetando al final de las líneas los géneros representados con el código:

```
plot_tree(physeq.betaproteo, color = "genotype_and_donor_status", shape = "Family",  
label.tips = "Genus", size = "abundance", plot.margin = 0.5, ladderize = TRUE)
```

Hay 3 géneros representados y 3 familias. La familia Alcaligenaceae con el género Sutterella representado, la familia Comamonadaceae con el género Acidovorax y Comamonas representado, y la familia Rhodocyclaceae, con ningún género asignado.

Los géneros están agrupados por su proximidad filogenética, por lo que se representan más juntos los que pertenecen a la misma familia.

En los organismos con el genotipo salvaje y estado saludable el género más abundante que tienen de la clase Betaproteobacteria parece ser el género Sutterella. En cuanto al genotipo susceptible y estado saludable (algo menos abundante que el anterior), pero que el género también mayor representado es Sutterella, seguido de Comamonas o Acidovorax, igual representados. El genotipo salvaje y con enfermedad de Parkinson parece el menos abundante, representado por el género Acidovorax y género no asignado de la familia Rhodocyclaceae.